Aalto University
School of Science

Tuomas Rintamäki

Demand and wind power scenarios for predicting power prices in a transmission-constrained system

Master's thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology in the Degree Programme in Engineering Physics and Mathematics.

Espoo, April 21, 2015

Supervisor: Professor Ahti Salo
Instructor: MSc (Eng) Jussi Uskola

Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

Aalto University
School of Science

ABSTRACT OF
MASTER'S THESIS

Author:	Tuomas Rintamäki		
Title:	Demand and wind power scenarios for predicting power prices in a transmission-constrained system		
Date:	April 21, 2015	**Pages:**	vi+67
Major subject:	Systems and Operations Research		
Minor subject:	Economics	**Code:**	Mat-2
Supervisor:	Professor Ahti Salo		
Instructor:	MSc (Eng) Jussi Uskola		

The capacity of intermittent wind power has increased substantially in Northern Europe in the past decade. Heavily temperature dependent demand is another weather-driven factor in the power market in this region. Because wind power and demand forecasts involve high uncertainties, we generate probabilistic scenarios with situation dependent and spatio-temporal information. Moreover, we analyze the correlation between the wind power and demand scenarios and their impact on cross-border capacities.

The scenarios are fed to a day-ahead market model to determine their marginal and joint impacts on price levels and volatility. The key finding is that the impacts of demand and wind power can amplify each other and cause low off-peak and high peak prices. Moreover, the difference between off-peak and peak prices in the Nordic countries is increased by cross-border exchange to Germany. We verify the statistical quality of the generated scenarios by showing that they have a low bias, and, thus, they can be taken into operative use to obtain useful information about positive and negative risks in the day-ahead market.

Keywords:	Power market modelling, wind power, electricity demand, cross-border capacity, nonparameric probabilistic forecast, scenario generation, forecast quality evaluation
Language:	English

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan koulutusohjelma

Aalto-yliopisto
Perustieteiden korkeakoulu

DIPLOMITYÖN
TIIVISTELMÄ

Tekijä: Tuomas Rintamäki

Työn nimi: Kulutus- ja tuulivoimaskenaariot siirtorajoitetun verkon sähkön hintaennustamisessa

Päiväys: April 21, 2015 **Sivumäärä:** vi+67

Pääaine: Systeemi- ja operaatiotutkimus

Sivuaine: Kansantaloustiede **Koodi:** Mat-2

Valvoja: Professori Ahti Salo

Ohjaaja: DI Jussi Uskola

Vaihteleva tuulivoimatuotanto on kasvanut merkittävästi Pohjois-Euroopassa viime vuosikymmenenä. Toinen sääriippuva tekijä alueen sähkömarkkinalla on lämpötilariippuva kulutus. Koska tuulivoima- ja kulutusennusteet ovat epävarmoja, tässä työssä rakennetaan todennäköisyyspohjaisia skenaarioita, jotka sisältävät tilanne-, paikka- ja aikariippuvaa tietoa. Lisäksi työssä analysoidaan näiden skenarioiden välistä korrelaatiota ja niiden vaikutusta maiden välisiin siirtokapasiteetteihin.

Skenaariot syötetään sähkön markkinahintamalliin, jonka avulla analysoidaan niiden yksittäis- ja yhteisvaikutusta sähkön hintaan ja volatiliteettiin. Tulokset osoittavat, että kulutuksen ja tuulivoiman vaikutukset voivat vahvistaa toisiaan ja aiheuttaa alhaisia yö- ja korkeita päivähintoja. Lisäksi yö- ja päivähintojen ero kasvaa usein Pohjoismaiden ja Saksan välisen sähkön siirron seurauksena. Tässä diplomityössä rakennetut skenaariot ovat tilastollisesti verraten harhattomia. Siten niitä voidaan hyödyntää operatiivisessa käytössä, jossa ne antavat hyödyllistä tietoa sähkömarkkinan riskeistä.

Asiasanat: Sähkömarkkinamallinnus, tuulivoima, sähkön kulutus, maiden välinen siirtokapasiteetti, parametriton todennäköisyysennuste, skenaarioluonti, ennusteen laadun arviointi

Kieli: Englanti

Acknowledgements

I would like to thank Jussi Uskola for the possibility to write this thesis and professor Ahti Salo and Tuomas Kervinen for their valuable feedback.

Moreover, I would like to thank my family and friends for supporting me in my studies.

Contents

1 Introduction

1.1 Northwestern European electricity markets

In early 2014, the most recent step in the deregulation process of the Nordic, Baltic and Central European electricity markets was taken by establishing the so-called Northwestern European Price Coupling (NWE) (Nord Pool Spot, 2014). The NWE region utilizes common day-ahead electricity price calculation in which accepted buy and sell bids and cross-border transmission flows are determined with the objective of maximizing social welfare.

If there are no transmission line congestions between different bidding areas, then a common market clearing price is found at the intersection point of the aggregated demand and supply curves. When transmission lines become congested, the NWE region is divided into price areas. For example, a deficit area imports from surplus areas until there is available transmission capacity but then the area price of the deficit area rises above the surplus areas. The area price is common to all participants in the area and the highest margin is earned by those generators whose marginal costs are lowest.

In practice, bids for the following day are submitted to the exchange by 12:00 noon (CET), after which the accepted bids and corresponding day-ahead prices and generation schedules are determined and published by 13:00 CET. At this point, the role of the transmission system operators (TSOs) is to optimize power flows inside the price areas given the generation schedules and to procure reserve power.

Due to transmission congestions inside a price area, generator failures and forecast errors of demand and intermittent renewable generation, intraday trading of all hours of the following

day takes place until 45-60 minutes before delivery. In real-time, the final balancing is done by the TSOs in regulation markets which are divided into primary, secondary and tertiary markets depending on the response time of the generators. Because of the integration of the electricity markets, also foreign reserves can balance the domestic power system. Moreover, there are financial products on weekly, monthly, and yearly day-ahead prices, for example, which can be traded before the delivery period starts. These products can be used for hedging the price risk or speculation. The timeline of the physical power markets is presented in Figure 1.

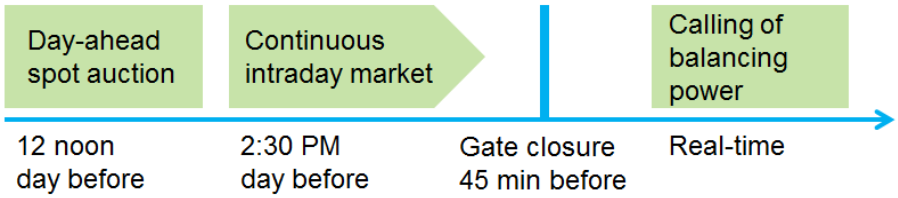


Figure 1: Timeline of physical power markets

Figure 2 shows that generation mix varies considerably in different Nordic countries and Germany. Sweden and Norway are hydro-dominant but Finland, Denmark, and Germany are thermal-based power systems. In general, hydro-dominant systems are characterised by high flexibility, which, however, can be locally limited if the reservoirs of the river system are small or if there are prolonged periods of wet hydrological conditions. In the Nordic system, spring and autumn floods often pose challenges to the operation of hydropower plants and require careful planning several weeks ahead. On the other hand, thermal-based systems have slower dynamics due to slow ramp up of coal condense plants. Therefore, thermal generation is best suited for base load generation, which does not need to change its output constantly. However, failures of single thermal plants can have a considerable impact on the operation of the system because their output is often several hundred megawatts. In these events, fast regulation can be done with gas turbine plants but these have often very high marginal costs.

Due to targets to reduce greenhouse gas emissions, intermittent renewable sources - wind and solar in particular - have recently gained large shares in Germany, Denmark and Sweden with installed capacities of 70300 MW, 4800 MW and 4500 MW in 2013, respectively (Bundesnetzagentur, 2014, The European Wind Energy Association, 2014). In Norway and Finland,

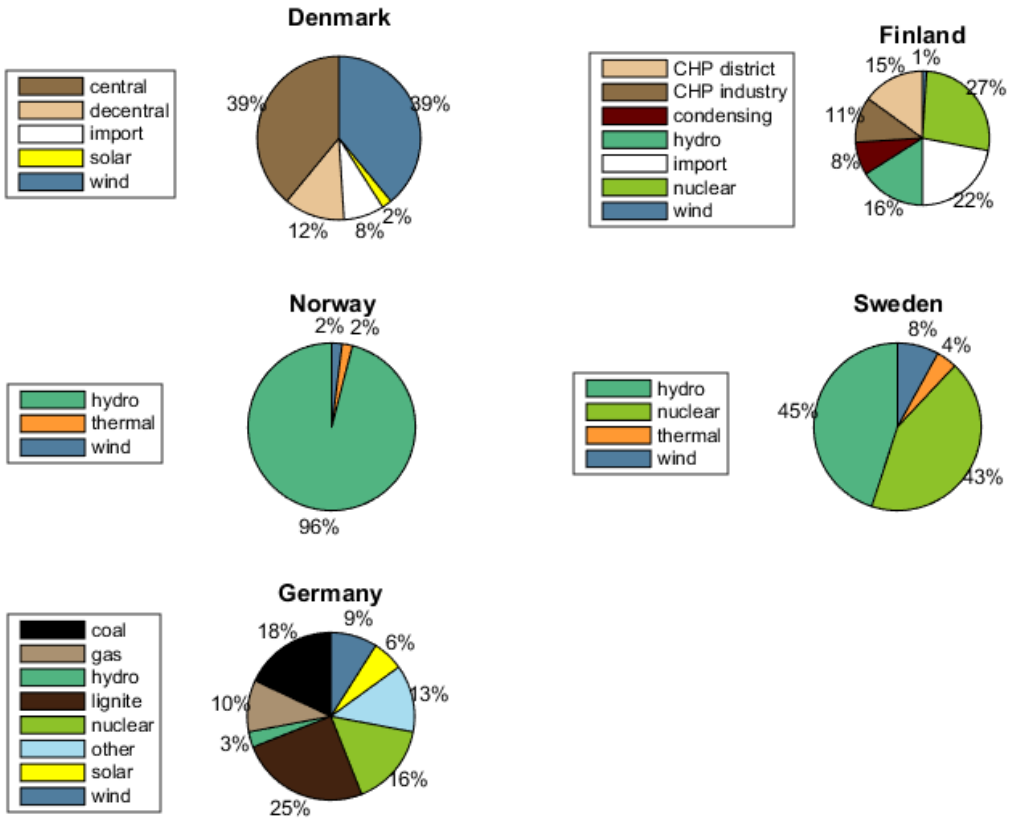


Figure 2: Generation mixes in the Nordic countries and Germany in 2014. Data source: Nordic TSOs and AG Energiebilanzen.

the share of wind power is still very small but since the fast growth of recent years is expected to continue, the share of wind power is forecasted to increase significantly in these countries, too. Moreover, large shares of nuclear power are present in Finland, Germany and Sweden.

Because of the economic downturn, electricity consumption has not grown in the Nordic countries and Germany for several years. Until 2020, consumption is forecasted to either decrease or stagnate as a result of low economic growth and improvements in energy efficiency Eurelectric (2013). However, some growth may come from electrification of transportation. In the long-term, climate change is forecasted to increase average temperatures, which will decrease heating-driven household consumption.

	DE	DK1	DK2	FI	NO	SE1	SE2	SE3	SE4
avg	37.7	35.3	36.4	37.9	32.3	34.1	34.1	34.5	35.3
stdev	11.8	15.7	9.9	9.7	8.6	8.9	8.9	9.3	9.5

Table 1: Average and standard deviation of daily prices in 2012-2014. NO is the average of all Norwegian areas.

Table 1 shows that the differences in generation mixes are reflected in the average and standard deviation of daily power prices in 2012-2014. Norway has the lowest average price due to its access to hydro and Danish wind power. Also, Norway has the lowest standard deviation of daily prices because of the flexibility of hydropower but the changes in yearly Norwegian prices can be high because of the differences between dry and wet hydrological years, whereas in thermal-based systems the yealy changes are relatively small. Swedish hydro power is mostly located in northern areas SE1 and SE2, whereas areas SE3 and SE4 with higher daily prices and standard deviation have higher demand and more nuclear and thermal production. Finland, on the other hand, has the highest average power price due to its lower generation capacity and the resulting dependency on imports. The average price for Germany is nearly as high as that for Finland, but its standard deviation is much higher due to the high share of intermittent renewables. The high standard deviation in Germany is only exceeded by DK1, where the share of wind power is highest.

In addition to the integration process, North Western European power system is undergoing several other changes that have an impact on the operation of the power system. Negligible marginal cost of renewables has decreased electricity spot prices (see Würzburg et al. (2013) for estimations of the impact) which has eroded the profitability of established utilities and led to the decommissioning of conventional power plants. However, the intermittency of renewables increases the volatility of day-ahead prices (Ketterer, 2014), which benefits the operators of flexible generation capacity who are capable of forecasting price peaks and drops. Large forecast errors of renewable generation increase volumes of intraday and balancing markets - +46 % year-on-year in Germany, for example (EEX, 2014) - and cause high deviations from the day-ahead

prices, which, in turn, allow flexible generation to make profit.

Because of increased cross-border transmission capacities, the volatility caused by renewables is expected to increase in the system. In particular, High Voltage Direct Current (HVDC) submarine power cables such as the Baltic Cable between Germany and Sweden and the Skagerrak 4 cable between Norway and Denmark have made flexible hydro power capacity available for continental Europe, on the one hand, and volatile wind power generation for the Nordic countries, on the other. To further integrate the European power system, the long-term network development plan of European Network of Transmission System Operators for Electricity (ENTSO-E, 2014) shows reinforcements to the German-Nordic corridor and new connections between Norway and wind-intensive Great Britain, for example. Moreover, studies such as OffshoreGrid (2014) have evaluated the feasibility of a meshed grid in the North and Baltic seas, concluding that such a grid would provide better access to offshore wind power, which is not utilized to large extent so far.

In Germany, in particular, the power system is moving from a classical power network with few centralised large power plants to a more dispersed one with a large number of small generation units. Additionally, new technologies such as smart metering devices, electric vehicles, storage, coupling of small-scale wind and thermal units are being developed, which further increases the complexity of the system. When the number of nodes increases and the dynamics of the system become faster, TSOs and distribution system operators (DSOs) face challenges in controlling the power flows in the network. When many distributed assets are located at the low voltage distribution level that was not, in the centralised framework, designed for absorbing generation. As a result, high renewable generation and its intermittency already set restrictions on the operation of the cross-border transmission lines in the day-ahead market.

1.2 Research objectives and assumptions

In an integrated power system with faster dynamics and higher uncertainty, there is a need for probabilistic forecasts. As the uncertainties are often weather related, the marginal and

joint impacts of probabilistic demand and wind power scenarios on cross-border transmission capacities and day-ahead power prices are analyzed in this thesis. In particular, we are interested in the possibility for high peak and low night time prices and their propagation in the Northwestern European electricity market. To simulate the day-ahead power prices, we develop a cost-minimization model which imitates the real price calculation algorithm.

The power price model includes the Nordic countries and Germany and exchange to neighbouring areas such as Estonia is following historical averages. For demand scenarios, all Nordic countries are covered but for wind power scenarios, only Denmark and Sweden are considered because they have the highest wind power capacity in the Nordics. Germany was not considered in the scenarios because it is represented only as an exogenous price forecast. These limitations are likely to have a small impact on the results.

Moreover, we assume perfect competition so that there is no strategic behaviour such as withholding of volumes to increase prices. Market data should be available transparently so that information such as generation cost parameters and transmission line capacities can be calibrated for the power price simulation model. Moreover, different price forecast inputs such as wind power and demand point forecasts are assumed to be given exogenously. For scenario generation, a database of earlier wind power and demand point forecasts and their realizations is required. The forecasts should not have systematic error as it can result into bias in the proposed methods.

The contribution of the thesis is to adopt recent nonparametric probabilistic methods and apply them to the Northwestern European energy market with realistic power market simulation. The statistical reliability of the methods is shown to be good, and, thus, the forecasting tools can be employed operatively in day-ahead spot and financial trading. Moreover, we discuss the correlation of the generated scenarios to provide insight into selecting the most likely scenario in each market situation.

Probabilistic forecasts have applications for forecasting intraday and balancing market volumes as they are dependent on wind power and demand forecast errors. Consequently, spatio-

temporal and situation dependent properties of wind power and demand are explored to find possibilities to apply the models in these markets. Only recently, modelling real-time markets have gained interest in the literature.

1.3 Structure of the thesis

This thesis is structured as follows: Section 2 reviews the literature on energy market modelling and scenario generation. Section 3 presents our day-ahead market model to which our demand and wind power scenarios from Section 4 are fed. The statistical validity of the scenarios is verified in Section 5 and power price simulation results presented in Section 6. Finally, Section 7 draws conclusions about the impact of the higher uncertainty in the market and discusses directions for future scenario generation models.

2 Earlier models of electricity markets

In this thesis, we employ bottom-up models of electricity markets which can consider endogenous price formation. One such approach is complementarity modelling, where a differentiable objective function such as generation cost function is minimized subject to differentiable constraint functions using Karush-Kuhn-Tucker (KKT) conditions. An overview of complementarity models can be found in Ruiz et al. (2013b) and the thorough treatise in Gabriel et al. (2013), who also consider the conditions for solving the optimization problem globally. For example, Leuthold et al. (2012) develop a large-scale perfect competition model of the European day-ahead electricity market covering transmission, variable demand, wind power, pump storage, and many other details.

Kunz (2013) imitates closely the German day-ahead and balancing power dispatch by presenting a sequential cost minimization model in which the generation schedules determined by a day-ahead market model ignoring the physical transmission network are fed into a separate congestion management model. Such models are successful only if the cost parameters of the generators are calibrated accurately. Many authors base their parameters on technology and fuel choices of the plants but this requires a great level of detail. On the other hand, Kännö (2013) establishes an aggregated cost function with an algorithm, which substracts net exchange and must-run supply such as nuclear and wind power from realised consumption and compares the resulting priced supply against realised market prices. Ruiz et al. (2013a) estimate cost parametes by inverse optimization in which the problem and its outcomes are known but some input parameters not.

Also, complementarity models can consider multiple agents and strategic behaviour. For example, Hobbs (2001) introduces a transmission grid owner and several generation companies

that seek to maximize their profits simultaneously. Utilizing information on the demand curve, the players can have an impact on the market outcome, and, thus, exert market power to raise equilibrium prices above perfectly competitive levels. Gabriel and Leuthold (2010) model the European power system as a Stackelberg game where a large generation company acts as a leader and the rest of the market as followers. They find that some European generation companies can act strategically and lift electricity prices in several countries. Consequently, our assumption of perfect competition may not be fully realistic. However, game-theoretic models are difficult to solve because they are highly non-convex and there can exists multiple equilibria (Ruiz et al., 2013b).

An alternative approach is empirical models which can employ observed market data directly. Kristiansen (2014) presents a regression model for Nord Pool system price with historical prices, hydro inflow and reservoir levels as independent variables. Zhou et al. (2006) extend the standard ARIMA time-series model with error correction, which models the forecast residuals with an additional ARIMA process. The drawback of these models is that they cannot capture the impact of transmission line congestions and they represent the impact of independent variables with a single coefficient when their impact is more dynamic. Also, these models are too conservative in the sense that they cannot predict the high volatility caused by renewable energy generation.

Stochastic models are often applied for pricing derivatives on electricity prices because they can generate a distribution of power prices. Geman and Roncoroni (2006) present a seasonal mean-reverting model with spikes that are drawn from a Poisson distribution. The model is calibrated to match the observed seasonality between winters and summers and spike occurrence. Vehviläinen and Pyykkönen (2005) employ a multifactor model where temperature and precipitation are modelled with stochastic processes which are inserted into stochastic models of demand and supply, which, in turn, are used for arriving at a spot price forecast. Such an approach suits for the Nordic countries, in particular, because hydropower has a high share and demand correlates strongly with temperature. However, also stochastic models cannot capture transmission line congestions and they are calibrated to historical observations, which

may not be representative for future. Also, the number of parameters increases fast when more fundamental factors are included.

Probabilistic power price forecasts can also be generated with quantile regression, where the conditional distribution of a dependent variable is modelled with given independent variable. Nowotarski and Weron (2014) benchmark parametric and semiparametric models and present a method called quantile regression averaging (QRA), where the point forecast of the best performing model is used as the independent variable for the actual spot price, which is the dependent variable. They find that simple quantile regression does not bring improvements over the best models, but QRA outperforms the earlier models.

Another class of proposed models is machine learning techniques such as neural networks, support vector machines, and hidden Markov models, which can be trained with a large set of fundamental input data (Weron, 2014). These models are flexible as they can handle complexity and non-linearity but according to Weron (2014), their performance is not yet tested thoroughly. Moreover, these techniques often become intractable and non-transparent for the user, and, thus, their usability in planning decisions is limited.

However, bottom-up models have the advantage that they can be used to explore the marginal impacts of fundamental factors and to derive theoretical results. Consequently, these models are widely employed to study the impacts of renewables on power systems. Using a bi-level model with wind and transmission grid investment decisions in the upper level and market clearing in the lower level, Baringo and Conejo (2011) show that investments in wind power require grid expansion. Spiecker et al. (2013) evaluate European transmission network expansion with projections of high wind power penetration until 2030. They find that stochastic wind and hydro power generation increase the value of grid expansion more than in the deterministic case and, thus, conclude that the grid extensions may increase the system flexibility and help integrating renewables. In particular, Jaehnert et al. (2013) find that the expansion of the North Western European HVDC network allows the hydro-dominant power systems in the Nordic countries to balance the variability of renewable generation in continental Europe by regulating hydro production accordingly.

These conclusions are confirmed in the light of recent market data by Mauritzen (2013) who finds that exports from Denmark to Norway and imports from Norway to Denmark are explained by high and low wind power generation, respectively. Effectively, Norway acts as a storage, where excess wind power is stored and later released at the time of low wind. In turn, Zugno et al. (2013) employ principal component analysis to obtain qualitative relationships between German wind power generation and cross-border flows but as such they cannot be used quantitively in forecasting.

On the other hand, Egerer et al. (2013) explore producer and consumer surplus in scenarios with different cross-border HVDC capacities and high renewable penetration. Producers in countries with low electricity prices, such as Norway and Sweden, gain from increased export possibilities, while producers in importing countries lose due to lower electricity prices. On the consumer side, the benefits are distributed the other way around through changes in prices. Therefore, the uneven distribution of the benefits of grid expansion may hinder its political acceptance. Moreover, Neuhoff et al. (2013) compare the operation of the European power system using zonal and nodal pricing and find that under zonal design transmission system operators have incentives to limit international transmission capacity to avoid domestic congestions.

Also, increased transmission network capacities help balance renewable-rich power systems. Farahmand and Doorman (2012) and Jaehnert and Doorman (2012) feed day-ahead generation schedules to a balancing market model where real-time imbalances and domestic transmission grid congestion occur. They find that the total need for re-dispatch is decreased by imbalance netting and the costs reduced by the exchange of balancing services through available transmission capacity in integrated European balancing markets. Based on the increasing volumes and profits in the balancing market, Boomsma et al. (2014) present a multi-stage stochastic program to develop a strategy to coordinate bidding to the day-ahead and balancing markets.

However, the previous authors use either realized balancing market volumes or generate them using a simple time-series model. Klæboe et al. (2015) benchmark these models for Norwegian balancing market volumes and prices and find their performance to be poor against a naive persistence model. They interpret the poor performance as a sign of the efficiency of

the Nordic electricity market, as balancing market is designed to handle only unexpected events and random fluctations in demand and supply.

However, because forecast errors of wind power and demand are not normally distributed, there can be occasions when it is very likely that the forecast is too low or high. These occassions give a trading signal for the day-ahead and balancing markets. Pinson et al. (2007a) benchmark the trading performance of a Dutch wind power farm with probabilistic wind power forecasts against point predictions and find the revenue of the farm to improve significantly. Consequently, generation of probabilistic scenarios of weather-dependent factors have the potential to improve planning decisions.

Different authors have proposed different parametric distributions for wind power forecast errors (Bludszuweit et al., 2008 and Hodge and Milligan, 2011), but generally the proposed distributions have fatter tails than the normal distribution. Pinson and Kariniotakis (2010) overcomes the discrepancy between different authors by employing empirical distributions for different forecast levels and look-ahead times to include time and situation dependent information. Moreover, Pinson et al. (2011) and Ma et al. (2013) present a wind power scenario generation framework where correlations between consecutive forecast errors are taken into account using inverse transformation method on multivariate normally distributed random numbers. Instead of temporal correlations, Grothe and Schnieders (2011) focus on spatial correlations of German wind power. Moreover, distributions can be generated for wind power ramps and generation cut-offs when wind speeds are too high or low.

Zhang et al. (2013) considers both spatial and temporal correlations of wind power production with high dimensional copula functions which give the conditional probability distribution of wind power production at certain location given recent observations in adjacent locations. In very high dimensions, single copula functions may become too complex, and, thus Valizadeh Zaghi and Lotfifard (2015) employ a novel vine copula method where the multidimensional copula function is split into bivariate components. Also, the method has the benefit that the distributions can be visualized. Alternatively, spatio-temporal wind power scenarios can be obtained

from weather ensemble forecasts (Nielsen et al., 2006) which, however, have limited temporal resolution in longer time periods.

Because of global warming, there is recent evidence of higher temperatures, precipitation and water inflows in Europe (Madsen et al., 2014). In the Nordic countries, in particular, electricity demand is driven by the need for heating. Consequently, also demand scenarios can provide a signal for trading in different electricity markets. A method similar to Pinson et al. (2011) and temperature ensemble forecasts has been applied to demand scenario generation, too (see Feng et al., 2013, and Taylor and Buizza, 2003). McSharry et al. (2005) utilize the temperature dependency and repetitive patterns of demand to develop probabilistic forecasts for time periods with no temperature forecast. They note that forecasting peak demand occurrence is particularly important for planning decisions.

In addition, different weather dependent components may have some mutual correlation structure. Qin et al. (2013) find a relatively strong positive correlation between demand and wind and solar power in a Chinese city and develop a method of correlated covariance matrices to obtain correlated samples. They note that considering these correlations are important for power system reliability assessment. Alternatively, the method of Pinson et al. (2011) can be augmented to handle wind power and demand at the same time. Furthermore, Baringo and Conejo (2013) employ k-means clustering to study how correlated wind and load in the Iberian Peninsula affect investment decisions.

With any of these simulation methods, a large number of scenarios need to be generated to obtain a proper sample. However, because our ultimate goal in this thesis is to generate price scenarios, the number of demand and wind power scenarios has to be limited to reduce computational burden. In the stochastic programming literature, different scenario reduction algorithms have been presented. Growe-Kuska et al. (2003) discusses optimal selection of scenarios by minimizing a distance between the reduced and original set of scenarios and presents algorithms for solving the problem. Also, they provide software to automatize the scenario selection.

For scenario generation, however, we found no commercial platform which could be utilised directly for our purposes, and, thus, the models need to be implemented in some programming language. For energy market modelling, there are many commercial candidates. Many references employ General Algebraic Modeling System (GAMS, 2015), which allows custom models to be specified and solved. Foley et al. (2010) review various other commercial energy modelling software, their features and main application areas. In general, these tools make it possible to map physical objects such as loads, transmission lines and generators to a particular optimization problem, and, thus, allow rapid development of models. Some software packages are flexible enough to offer tools for operative planning and investment analysis.

3 The day-ahead market model

3.1 Notation

Sets and indices

$n, k \in \mathrm{N}$	nodes
$t \in \mathrm{T}$	time horizon
$\ell \in \mathrm{L}$	transmission lines

Parameters

$d_{n,t}$	demand at node n during period t (MW)
$s_{n,t}(\cdot)$	supply curve function at node n during period t in the day-ahead market (€/MW)
$Y_{\ell,n}$	the direction of the transmission line ℓ at node n ($\{-1,0,+1\}$)
$NTC^{+}_{\ell,t}$	net transfer capacity to positive direction on line ℓ during period t (MW)
$NTC^{-}_{\ell,t}$	net transfer capacity to negative direction on line ℓ during period t (MW)

Signed variables

$\lambda_{n,t}$	price at node n during period t in the day-ahead market (€/MW)

Positive variables

$g_{n,t}$	generation at node n during period t (MW)
$tf_{\ell,t}$	transmission flow in line ℓ during period t in the day-ahead market (MW)
$\overline{\mu_{\ell,t}}$	dual variable for positive transmission flow in line ℓ during period t (MW)
$\underline{\mu_{\ell,t}}$	dual variable for negative transmission flow in line ℓ during period t (MW)

3.2 Model formulation

We consider a simplified network for the Nordic countries and Germany, where each price area is represented as a node (see Figure 3). Following the price calculation principles in Nord Pool Spot (2015), the objective function in Eq. (1) is to minimize the hourly total costs of generation $\sum_n \int_0^{g_{n,t}} s_{n,t}(g_{n,t})$ with respect to generation decisions $g_{n,t}$ and transmission flows $tf_{\ell,t}$ for given demand $d_{n,t}$. This objective is equivalent to maximizing overall social welfare, when demand is fixed. Eq. (2) requires that demand matches supply and possible cross-border flows in each node, where the flows are constrained by exogenously given net transfer capacities $NTC^{+}_{\ell,t}$ and $NTC^{-}_{\ell,t}$ in Eqs. (3)-(4). Moreover, the direction of the transmission line is determined by the element $Y_{\ell,n} \in \{-1,0,+1\}$ of an incidence matrix Y. As a result, export from node n to k in line ℓ is interpreted as demand at node n and generation at node k. The optimization is repeated for each period t sequentially.

$$\min_{g_{n,t} tf_{\ell,t}} \sum_n \int_0^{g_{n,t}} s_{n,t}(x)dx \tag{1}$$

s.t.

$$d_{n,t} + \sum_\ell Y_{\ell,n} tf_{\ell,t} = g_{n,t} \qquad \lambda_{n,t} \text{ free} \qquad \forall n \tag{2}$$

$$tf_{\ell,t} \leq NTC^{+}_{\ell,t} \qquad \overline{\mu_{\ell,t}} \geq 0 \qquad \forall \ell \tag{3}$$

$$-tf_{\ell,t} \geq NTC^{-}_{\ell,t} \qquad \underline{\mu_{\ell,t}} \geq 0 \qquad \forall \ell \tag{4}$$

Prices in each area are given by the dual variables $\lambda_{n,t}$ of the balance equations 2. Usually, prices are positive, but they can go negative if there is must-run generation, which is bidded at negative price. The dual variables $\overline{\mu_{\ell,t}}$ and $\underline{\mu_{\ell,t}}$ of Eqs. (3) and (4), respectively, give the improvement in the objective function if the transmission capacity is relaxed.

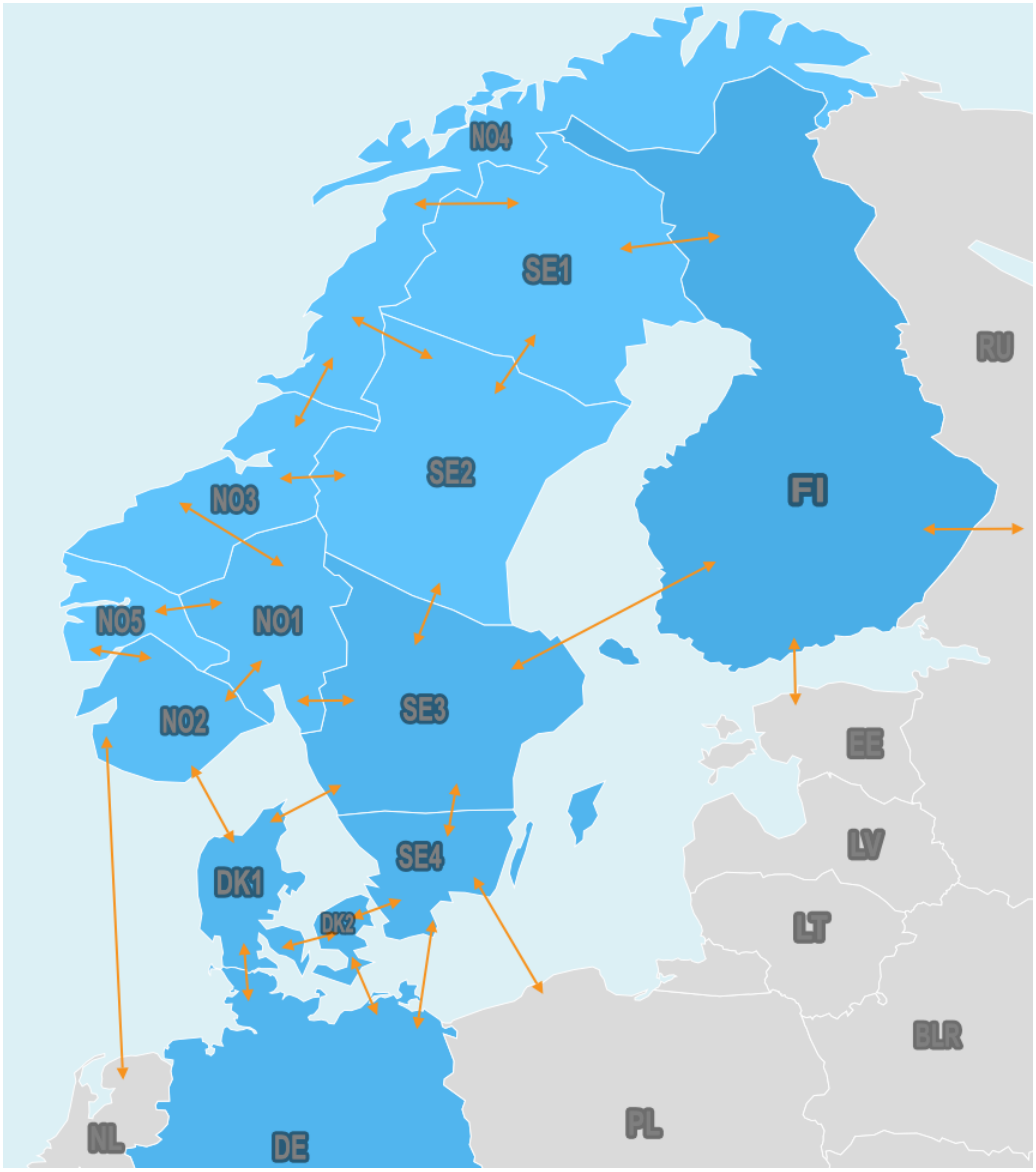


Figure 3: Bidding areas in the day-ahead market, i.e., the nodes of the model. Image source: Nord Pool

3.3 Calibration

To calibrate the model, only cost functions $s_{n,t}(\cdot)$ need to be specified as realised consumption, production and transmission capacities can be obtained from Nord Pool, EEX, different national TSOs or ENTSO-E. One alternative is to study individual power plants and determinine their fixed and variable costs (see e.g. Leuthold et al., 2012) but Kännö (2013) divides aggregate supply to price independent and dependent components: the former includes wind, CHP and nuclear power, for example, and the latter hydro and condensing power. Price independent supply is assumed to be bidded at negligible price and it is nearly constant (nuclear power) or it can be forecasted based on weather forecasts (wind power and CHP). Consequently, price dependent supply, i.e., the part of the supply curve with positive price can be calibrated by substracting price independent supply from consumption and net exchange. Using this methodology, we arrive at an aggregated cost function for each node in the model.

The functions $s_{n,t}(\cdot)$ represent merit-order for the areas. Figure 4 shows an example of the demand and supply curve for the whole Nord Pool area from which the Nord Pool system price, i.e., the intersection point, can be computed. The price independent supply extends up to 25 GW as the price of supply is less than or equal to 0 €/MWh. Then the price of supply gradually increases and close to 50 GW, the supply functions is nearly vertical, which means that there is no extra supply available. We can also read that a large part of the price dependent supply is hydro power priced at approximately 30 €/MWh. Because the supply curves do not, in general, have a functional form, they need to be handled as a piecewise linear function or approximated with a high-order polynomial, for example.

To use the model for price forecasting, different inputs need to be forecasted. Often, consumption can be forecasted using statistical models because of repetitive patterns in consumer behaviour (e.g. working day and weekend variation, holidays) and strong temperature dependency (see Huovila, 2003). However, forecasting of supply is more difficult due to adaptation to changing conditions and continuous profit maximization of the players and unexcepted failures of the generation units. Consequently, price independent and dependent supply need to be

updated from time to time. In addition, net transfer capacities between the price areas and possible power plant outages need to be specified according to the reported capacities in urgent market messages (UMMs). However, obtaining these inputs is outside the scope of the thesis and, thus, they are given exogenously.

Contrary to (possibly extremely nonlinear) game-theoretic models with hundreds or thousands of nodes, the model presented here is computationally light. Thus, it can be used to simulate relatively long time periods with hourly resolution. Moreover, a small number of nodes limits the number of parameters that need to be estimated and maintained. To simplify the model further, conditional factors such as block bids that span multiple time periods as well as rarely applied technical constraints such as transmission flow ramping and area-specific sum of export/import capacity limits have been neglected.

We assume that there is no strategic behaviour. If such behaviour were to occur, the relevant information could be to some extent incorporated into the cost functions as they calibrated to observed market data. However, a drawback of our approach is that the behaviour may change in future. Moreover, demand is fixed as the price elasticity of demand is assumed to be very low even with high prices (see the demand curve in Figure 4).

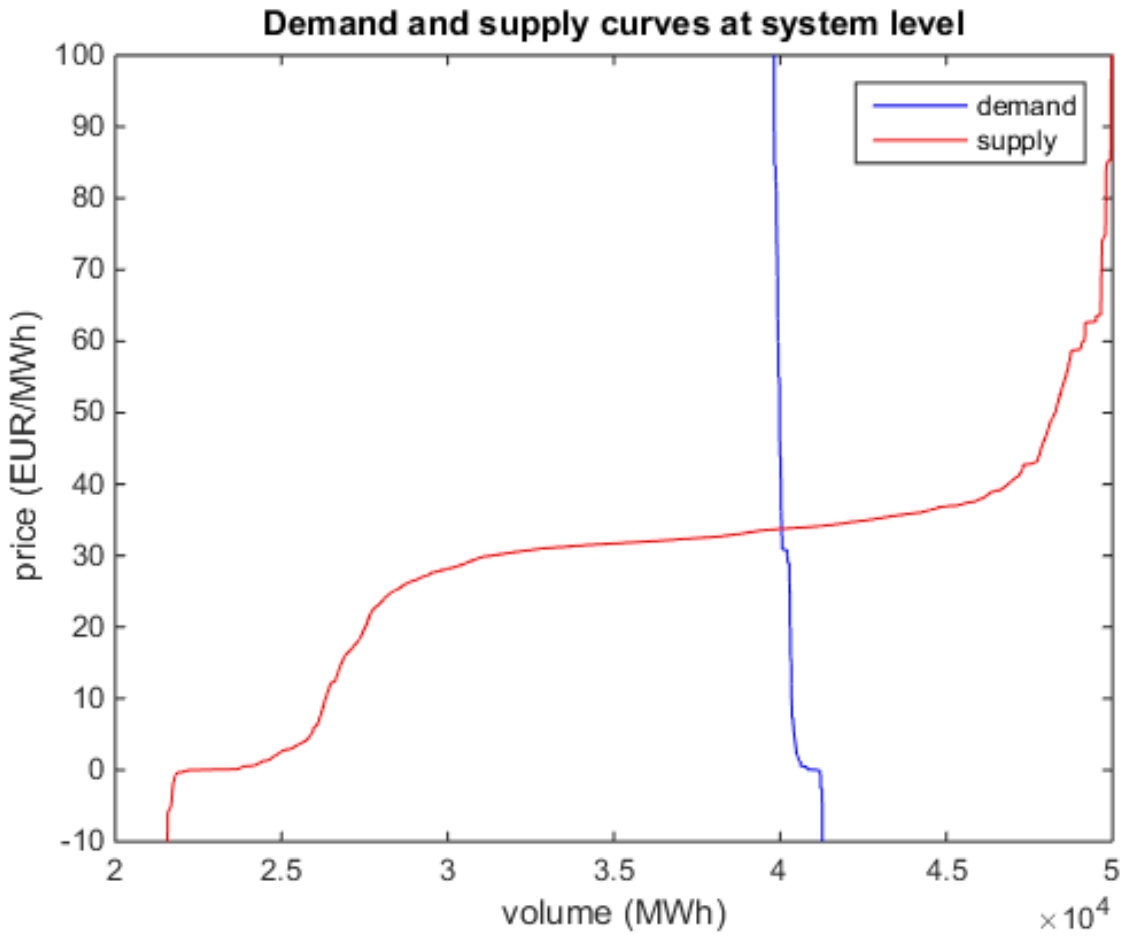


Figure 4: An example of demand and supply curves at Nord Pool system level on 6 October 2014, 18 CET

4 Demand and wind power scenario generation

In this section, the inverse transformation method is reviewed. This methodology is used to generate scenarios from our short- and medium-term models for wind power and demand. A large number of scenarios is generated, so we present a method for reducing their number for efficient power price simulation. Finally, the correlation between wind power and demand and their forecast errors is discussed and their impact on cross-border capacities analyzed.

4.1 Quantile forecasts and inverse transformation method

The standard regression analysis provides a estimate z of the conditional mean $\mathrm{E}[Y|X=x]$ of a dependent variable Y given the independent variable $X=x$ by minimizing the squared error

$$\mathrm{E}[Y|X=x] = \arg\min_{z} \mathrm{E}[(Y-z)^2|X=x]. \tag{5}$$

However, the conditional mean does not give any information about uncertainty of the estimate, which is described quantiles. To this end, let f_{t+k} denote the probability density function (pdf) and F_{t+k} the cumulative distribution function (cdf) of a random variable P_{t+k}. According to Pinson and Kariniotakis (2010), the quantile q^{α}_{t+k} of P_{t+k} is defined as the value x such that

$$\Pr(P_{t+k} \leq x) = F_{t+k}(x) = \alpha. \tag{6}$$

In other words,

$$q^{\alpha}_{t+k} = F^{-1}_{t+k}(\alpha). \tag{7}$$

The quantile forecast issued at time t for look-ahead time k is then denoted by $\hat{q}^{\alpha}_{t+k|t}$ and a set of n quantile forecasts is given by

$$\hat{f}_{t+k|t} = \left\{ \hat{q}^{\alpha_i}_{t+k|t} | 0 \leq \alpha_1 < ... < \alpha_i < ... < \alpha_n \leq 1 \right\}, \tag{8}$$

where α_i are selected nominal proportions spread on the unit interval. Following Pinson and Kariniotakis (2010), these probabilistic forecasts are referred to as predictive distributions. A standard method for obtaining the predictive distribution is quantile regression

$$q^{\alpha}_{t+k} = \arg\min_{z} \mathrm{E}[L_{\alpha}(Y,z)|X = x], \tag{9}$$

where

$$L_{\alpha}(y,z) = \begin{cases} \alpha|y-z|, & y > z \\ (1-\alpha)|y-z|, & y \leq z. \end{cases} \tag{10}$$

A more flexible framework is presented in Pinson et al. (2011) and Ma et al. (2013), who employ the inverse transform method for simulating paths of wind power forecast errors and arriving at a predictive distribution of wind power generation. In this method, the random variable P_{t+k} is sampled as follows

$$P_{t+k} = F^{-1}_{t+k}(U), \quad U \sim \text{ Unif}[0,1]. \tag{11}$$

Because the cdf of a random number Z_{t+k} sampled from a standard normal distribution with zero mean and unit standard deviation is uniformly distributed over $[0,1]$, Eq. (11) becomes

$$\begin{aligned} P_{t+k} &= F^{-1}_{t+k}(\Phi(Z_{t+k})), \text{ where} \\ \Phi(Z_{t+k}) &= \int_{-\infty}^{Z_{t+k}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \end{aligned} \tag{12}$$

The transformation is illustrated in Figure 5. The advantage of this transformation is that random numbers can be sampled from any distribution by generating uniformly distributed random numbers. An alternative sampling methodology is rejection sampling (Jordan, 2010).

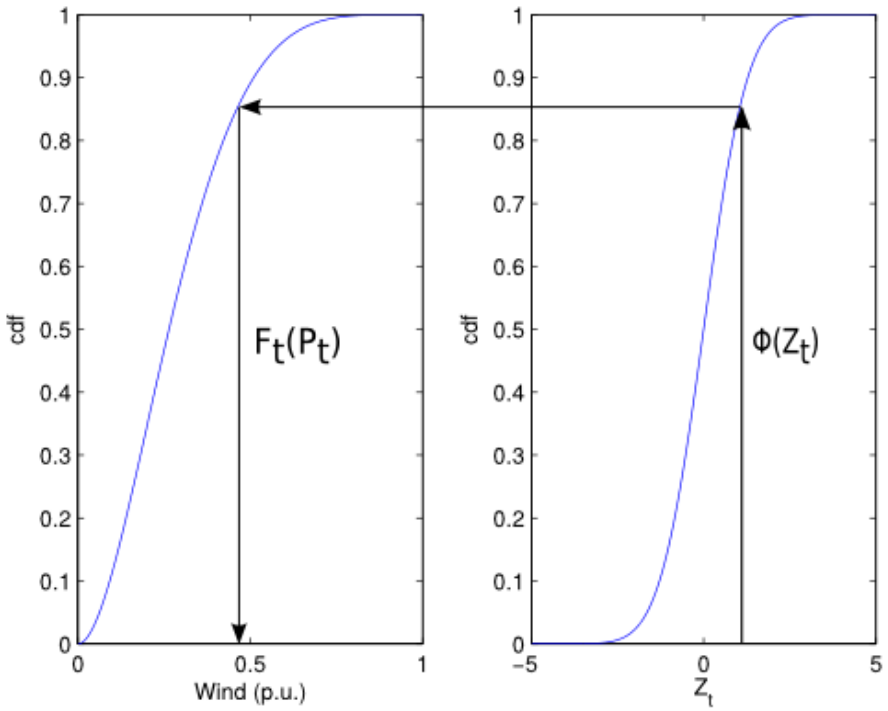


Figure 5: An illustration of the inverse transformation method. The value of the random variable Z_t is mapped to a wind power forecast error P_t

4.2 Empirical distributions and normal multivariate random numbers

To use the inverse transformation method, one needs to specify the inverse cumulative distribution function F_{t+k}^{-1} and random number Z_{t+k}. In this thesis, F_{t+k} and F_{t+k}^{-1} are nonparametric empirical distributions. To illustrate, the procedure for obtaining them for forecast errors is as follows. First, a set of old point forecasts $f_{t+k|t}$ issued at time $t = 1 \ldots n$ for each *look-ahead time* $k = 1 \ldots N$ is gathered and divided into b bins $i = 1 \ldots b$, which are called *forecast conditions* $c_{i,k}$. The set of forecast errors in the forecast condition bin $c_{i,k}$ is given by $E^{c_{i,k}} = \left\{ f_{t+k|t} - r_{t+k} |\ c_i \leq f_{t+k|t} < c_{i+1}, t = 1 \ldots n \right\}$, where r_{t+k} is the realisation at time $t+k$. The corresponding empirical cumulative distribution function (ecdf) is given by

$$\hat{F}^{c_{i,k}}(x) = \frac{1}{m} \sum_{e \in E^{c_{i,k}}} I(e \leq x), \tag{13}$$

where e is a forecast error in the set $E^{c_{i,k}}$ and m the number of observed errors and $I(\cdot)$ the indicator function, which gives 1 if its argument is true and 0 otherwise. If the ecdf needs to be smoothed for differentiability, then a kernel function $K : \mathbb{R} \to \mathbb{R}, \int_{-\infty}^{\infty} K = 1$ with a smoothing parameter h can applied, which gives

$$\hat{F}^{c_{i,k}}(x) = \frac{1}{mh} \sum_{e_i \in E^{c_{i,k}}} K\left(\frac{x - e_i}{h} \right). \tag{14}$$

A popular choice is the Gaussian kernel smoother $K(x, x_i) = \exp\left(-\frac{(x-x_i)^2}{2h^2} \right)$, which is readily available in numerical software such as Matlab.

The significance of having distributions for each forecast condition and look-ahead $c_{i,k}$ is shown in Figure 6. If the forecast is very low (subfigure $c_{1,1}$), then it is statistically very unlikely that the forecast error would be negative, i.e., the realisation would be even lower than the forecast. On the other hand, if the forecast is very high ($c_{15,1}$), then the forecast is almost always too high. In both cases, the absolute forecast error increases with time and at day 9 ($c_{1,9}$ and $c_{15,9}$) the point forecast is practically useless as the forecast error can be as high as 3000 MW in this example.

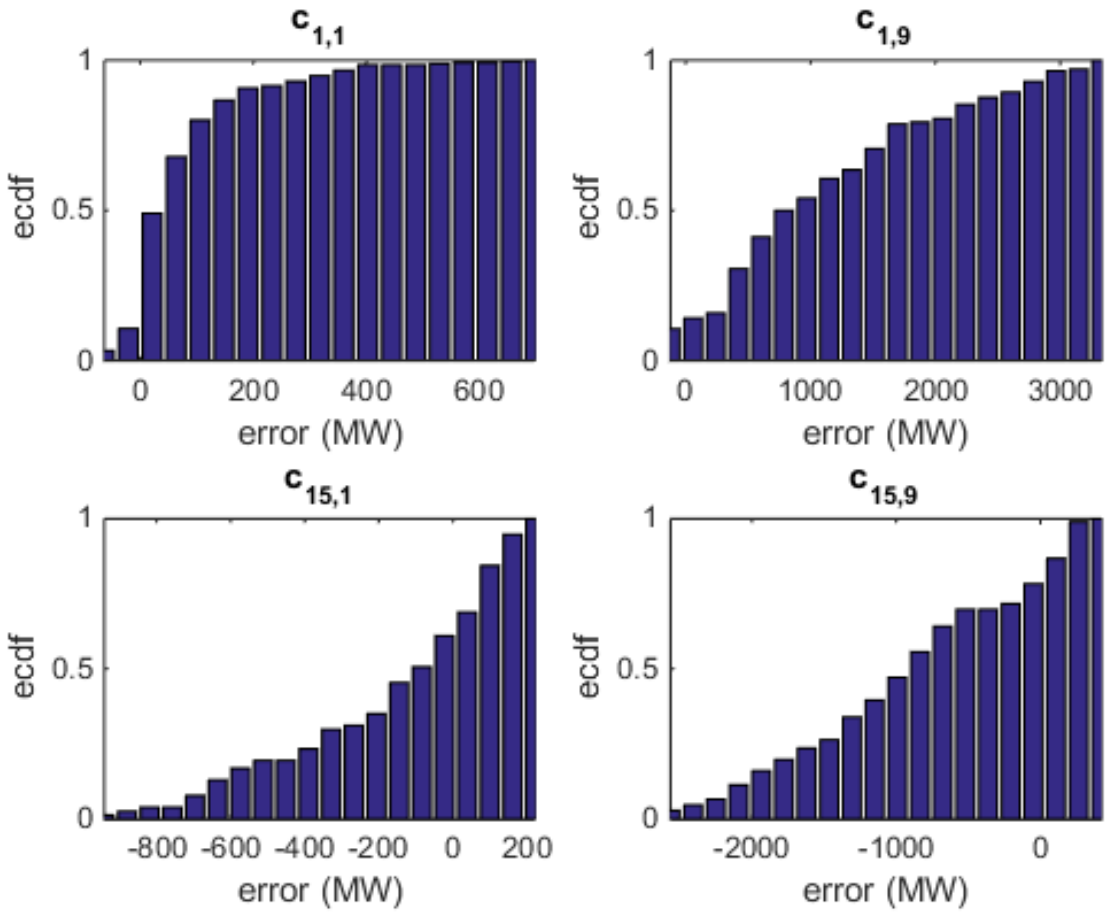


Figure 6: Empirical cumulative distribution functions for different forecast conditions and look-ahead times

As we wish to generate a vector of forecast errors $\mathbf{P} = [P_1, P_2, ..., P_N]$, where N is the end of the forecast horizon, a random vector $\mathbf{Z} = [Z_1, ...Z_N]$ needs to be generated. To incorporate temporal correlation in the forecast errors, the consecutive random numbers Z_t and Z_{t-1} need to be correlated. Following Pinson et al. (2011) and Ma et al. (2013), we employ the simplest alternative for the joint distribution of $\mathbf{Z}$ by assuming that it follows the multivariate normal distribution $\mathbf{Z} \sim N(\mu, \Sigma)$, where μ is a N-dimensional vector of zeros and the covariance Σ a positive semi-definite $N \times N$ matrix.

Random numbers with such covariance structure can be generated with the help of Cholesky decomposition, in which the problem is to find a lower triangular matrix L such that $\Sigma = LL^T$, where L^T denotes the transpose of L. Then, using uncorrelated random variables with zero mean and unit variance, we obtain $Z = LX$ as $\mathrm{E}[ZZ^T] = \mathrm{E}[(LX)(LX)^T] = \mathrm{E}[LXX^TL^T] = L\mathrm{E}[XX^T]L^T = LIL^T = LL^T = \Sigma$. Cholesky decomposition exists only for positive semidefinite matrices because $x^T\Sigma^T x = x^T L^T Lx = (Lx)^T(Lx) = ||Lx||^2 \geq 0$.

The covariance matrix Σ can be estimated from historical forecast errors by applying the inverse transformation method backwards. Each forecast error $e_{t+k|t}$ of a forecast issued at time t is categorized according to its look-ahead time k and the associated forecast $f_{t+k|t}$ to a forecast condition $c_{i,k}$ and transformed to uniform variable u_k using the estimated empirical cumulative distribution function $\hat{F}^{c_{i,k}}(x)$. The variable u_k is then transformed to a normally distributed variable Z_k using the inverse cdf of the standard normal distribution. By repeating this process, the forecast errors are compiled into vector $\mathbf{Z_t}$ for each forecast time. The sample covariance matrix is then given by

$$\mathbf{Q} = \frac{1}{n-1}\sum_{t=1}^{n}(\mathbf{Z}_t - \overline{\mathbf{Z}})(\mathbf{Z}_t - \overline{\mathbf{Z}})^T, \tag{15}$$

where

$$\overline{\mathbf{Z}} = \begin{bmatrix} \bar{Z}_1 \\ \vdots \\ \bar{Z}_N \end{bmatrix} = \frac{1}{n}\sum_{t=1}^{n}\mathbf{Z}_t. \tag{16}$$

However, this empirical estimate can be singular, which prevents the use of Cholesky decomposition. Therefore, following Ma et al. (2013), we employ an exponential covariance function

with parameter ε controlling for the strength of the correlation

$$\text{cov}(Z_i, Z_j) = \exp\left(-\frac{|i-j|}{\varepsilon}\right) \quad i, j = 1...N. \tag{17}$$

The correlation between consecutive forecast errors can be studied by looking at the their first differences $d_t = e_t - e_{t-1}$. In the following subsections, ε is calibrated using the following optimization problem

$$I_\varepsilon = \min_{\varepsilon} |\text{ecdf}(d) - \text{ecdf}(d')|, \tag{18}$$

where $\text{ecdf}(d)$ and $\text{ecdf}(d')$ are the empirical cumulative distribution functions of the first differences of the observed and simulated forecast errors, respectively. Ma et al. (2013) note that the optimal value of ε changes with time and Pinson et al. (2011) employ a adaptive scheme for updating the covariance structure. However, given our higher level target of generating spot price scenarios, slightly suboptimal values of ε will not be of great importance, and, thus, it will be sufficient to re-estimate this parameter on a regular basis.

4.3 Scenario reduction

The inverse transform method is used to sample a large number of scenarios. Here, we employ a very simple scenario reduction scheme, which is needed to simulate power prices within a time limit set by operative use. The generated scenarios of a random variable such as wind power are averaged over the forecast horizon and an empirical cumulative distribution function is generated from these averages. In the case of forecast errors, in particular, the distribution can be skewed if the associated forecast is very high or low for the whole horizon.

From this distribution, the average values corresponding to different predetermined quantiles are obtained. Then, for each quantile, the set of scenarios is iterated until a path is found whose average value matches that of the quantile within certain threshold δ. The value of δ depends on the random variable but it should be kept as small as possible so that a scenario path with significantly lower or higher average does not get selected. As a result of this process, a desired number of scenarios is obtained selected, which, however, may not be the most probable but

they will give the sensitivity and the absolute level of power prices even in relatively unlikely events that may be important for risk analysis.

4.4 Wind power scenarios

Wind power scenarios are generated in two phases. First, we consider the point forecast for wind power to be accurate enough for the first $N = 9$ days $= 216$ hours. Within this horizon, the method of Pinson et al. (2011) and Ma et al. (2013) is applied to simulate forecast errors, which are added to the point forecast. In our model, the forecasts for hours 1-24, 25-48, and so on are categorized to look-ahead times $k = 1, 2, \ldots$, respectively, as the forecast errors are assumed to stay nearly constant within each day. Moreover, there are $b = 15$ evenly distributed forecast condition bins between the minimum and maximum forecast and each ecdf is divided into 30 bins. We note that these parameters are arbitrary and their optimal value depends on time. As the forecast levels and forecast errors depend on the time of the year, the historical forecast errors used in the estimation of empirical cumulative distribution functions should be selected from the same time of the year as the forecast is made for. The procedure for generating scenarios follows the discussion in Sections 4.1 and 4.2 and it is outlined in Algorithm 1.

The results for calibrating the parameter ε for DK1 and DK2 areas using Eq. (18) are presented in Figure 7. The objective function value improves considerably when ε increases from the initial value. After $\varepsilon = 150$, however, the objective becomes worse or stays nearly the same. The optimal value for both areas is roughly $\varepsilon = 100$, which gives the best temporal correlation. The optimal value of ε indicates that there is strong positive correlation between consecutive and past forecast errors.

The second phase is to generate forecast for day 10 onwards up to 28 days. In this horizon, the point forecast is the wind power normal, i.e., the average of historical hourly wind power generation during the same time of the year. The same procedure in Algorithm 1 can applied here but instead of sampling forecast errors, realised wind power production should be used.

Algorithm 1: Procedure for generating short-term wind power scenarios

1. The empirical cumulative distribution functions of observed forecast errors $\hat{F}^{c_{i,k}}(x)$ are computed for each combination of forecast condition and look-ahead time $c_{i,k}$ and stored.

2. $D = 1000$ realizations of N-dimensional random vectors $\mathbf{Z} \sim N(\mu, \Sigma)$ are generated using the Cholesky decomposition. The element for kth look-ahead time in dth scenario is denoted with $Z_{k,d}$.

3. For each look-ahead time in the most recent wind power point forecast $f_{t+k|t}$, determine to which forecast condition bin $c_{i,k}$ the forecast belongs. The corresponding ecdf $\hat{F}^{c_{i,k}}(x)$ is adopted.

4. The inverse transformation method in Eq. (12) is applied with $[\hat{F}^{c_{i,k}}]^{-1}$ and $Z_{k,d}$ to simulate a forecast error for each look-ahead time in each scenario, i.e., the set of errors $E = \{[\hat{F}^{c_{i,k}}]^{-1}(\Phi(Z_{k,d})) \,|\, d = 1 \dots D\}$.

5. D scenarios are reduced using the method described in Section 4.3. The selected paths are added to the point forecast $f_{t+k|t}$ to obtain a set of quantile forecasts in Eq. (8).

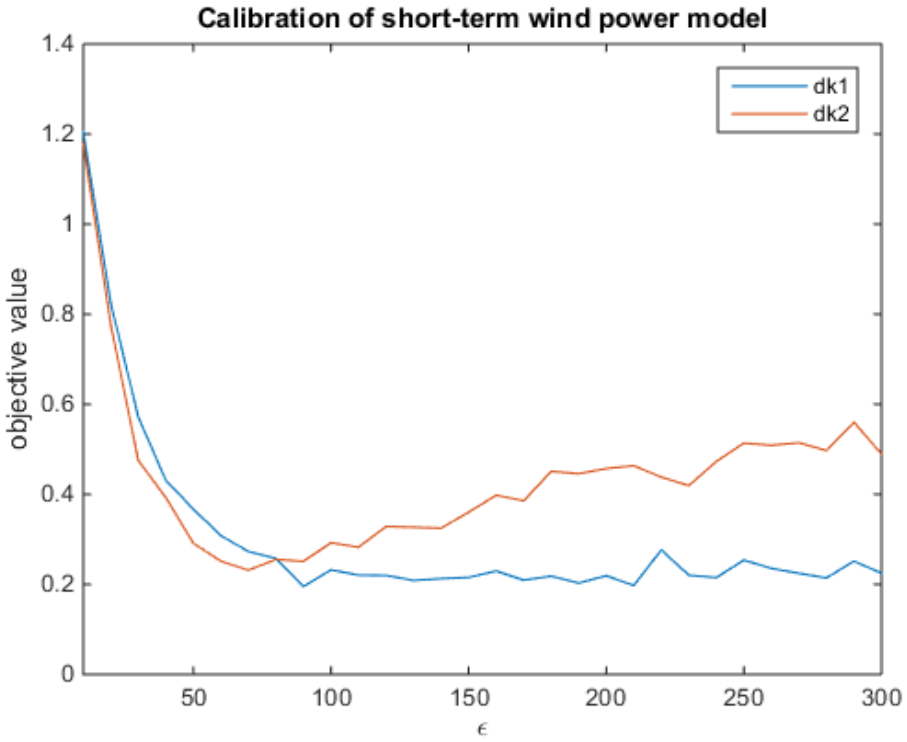


Figure 7: Calibration of the ε parameter for DK1 and DK2.

However, we prefer the copula model in Zhang et al. (2013) as the wind power realisations in the Nordic countries exhibit a strong positive spatial correlation in addition to temporal correlation (see Table 2). Next, we briefly review the copula theory.

According to the Sklar's theorem, the joint cumulative distribution of two d-dimensional random variables X and Y can be modelled with their individual cumulative distributions F_X and F_Y and a copula function $C : [0,1]^d \times [0,1]^d \to [0,1]$

$$F_{XY}(x,y) = C(F_X(x), F_Y(y)). \tag{19}$$

When C, F_X and F_Y are differentiable, we obtain the probability density of the joint distribution

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} = c(F_X(x), F_Y(y)) f_X(x) f_Y(y), \tag{20}$$

where f_X and f_Y are the probability density functions of the marginals and $c(\cdot)$ the density of the copula function.

	$DK1_t$	$DK1_{t-1}$	$DK2_t$	$DK2_{t-1}$	SE_t	SE_{t-1}
$DK1_t$	1	0.99	0.83	0.81	0.64	0.63
$DK1_{t-1}$		1	0.83	0.83	0.66	0.64
$DK2_t$			1	0.99	0.56	0.55
$DK2_{t-1}$				1	0.57	0.56
SE_t					1	0.99
SE_{t-1}						1

Table 2: Spatio-temporal correlations of Nordic wind power in 2013-2014

Using the definition of conditional probability, we have

$$f(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = c(F_X(x), F_Y(y)) f_X(x). \tag{21}$$

Like Zhang et al. (2013), we let $X^t = [x_1^t, ..., x_n^t]$ and $X^{t-1} = [x_1^{t-1}, ..., x_n^{t-1}]$ denote the wind power production of n locations during time periods t and $t-1$, and $U = [F(x_1^t), ..., F(x_n^t)]$ and $V = [F(x_1^{t-1}), ..., F(x_n^{t-1})]$ their transformed margins. Using Eq. (21), the conditional probability distribution function with spatio-temporal dependencies can be expressed with

$$f(X^t|X^{t-1}) = \frac{f(X^t, X^{t-1})}{f(X^{t-1})} \tag{22}$$

$$= \frac{c(U,V)\prod_{i=1}^{n} f(x_i^t)\prod_{j=1}^{n} f(x_j^{t-1})}{c(V)\prod_{j=1}^{n} f(x_j^{t-1})} \tag{23}$$

$$= \frac{c(U,V)}{c(V)}\prod_{i=1}^{n} f(x_i^t). \tag{24}$$

From Eq. (24), Zhang et al. (2013) derive the following formula for the probability distribution function of wind power production at location i at time t

$$f(x_i^t|x_1^t, \ldots x_{i-1}^t, X^{t-1}) = \frac{c(F(x_1^t), \ldots, F(x_{i-1}^t), F(x_i^t), V)}{c(F(x_1^t), \ldots, F(x_{i-1}^t), V)} f(x_i^t). \tag{25}$$

Note that if $i = 1$ above, then terms $x_1^t, \ldots, x_{i-1}^t$ and $F(x_1^t), \ldots, F(x_{i-1}^t)$ drop out in the conditional

probability and copula functions, respectively, and we have

$$f(x_i^t|X^{t-1}) = \frac{c(F(x_i^t),V)}{c(V)} f(x_i^t). \tag{26}$$

In Eq. (25), the copula functions can be very high-dimensional, in particular if observations earlier than $t-1$ are accounted for. To reduce the dimensions, we drop the vector of earlier observations X^{t-1} but rather let x_i^t be temporally correlated by using correlated multivariate normal random numbers, which are, similar to the short-term model, calibrated using Eq. (18). Then, Eq. (25) becomes

$$f(x_i^t|x_1^t,\ldots x_{i-1}^t) = \frac{c(F(x_1^t),\ldots,F(x_{i-1}^t),F(x_i^t))}{c(F(x_1^t),\ldots,F(x_{i-1}^t))} f(x_i^t). \tag{27}$$

For $i=1$, we have only its own probability density function $f(x_1^t)$. Therefore, wind power for the first area can be sampled by applying the inverse transformation method to its empirical cumulative distribution function $F(x_1)$, which is estimated from realised generation. The drawback of this simplification is that wind power for area $i=1$ is not spatially correlated with other areas anymore. However, low pressure systems often move from West to East in the Nordics, and, thus, we can use the westernmost area, i.e., DK1 as the first area.

Here, the copula functions c need to be estimated from realised data, which should be from the same time of the year as the forecast is made for because the level of wind power generation increases from summer to winter and decreases again in the spring. As the geographical area in question is large, we can conclude that wind power data is asymmetric so that $C(u,v) \neq C(v,u)$, and, thus, a simple symmetric copula function such as Gaussian copula may not be adequate. Moreover, the different marginal distributions do not need to follow the same distribution making the problem even more complex. Therefore, similar to short-term wind power data, empirical copula functions are preferred. Such an empirical copula density for 1-dimensional random variables X and Y with a grid division parameter h is defined as follows (Charpentier et al., 2006)

$$\hat{c}(u,v) = \frac{1}{Nh^2} \sum_{i=1}^{N} K\left(\frac{u-F_X(X_i)}{h}, \frac{v-F_Y(Y_i)}{h}\right), \tag{28}$$

where K is a bivariate kernel function $K:\mathbb{R}^2 \to \mathbb{R}$ such as the multiplicative kernel $K(u,v) = K_1(u)\cdot K_2(v)$, where K_1 and K_2 are univariate kernels such as the Gaussian kernel. Together

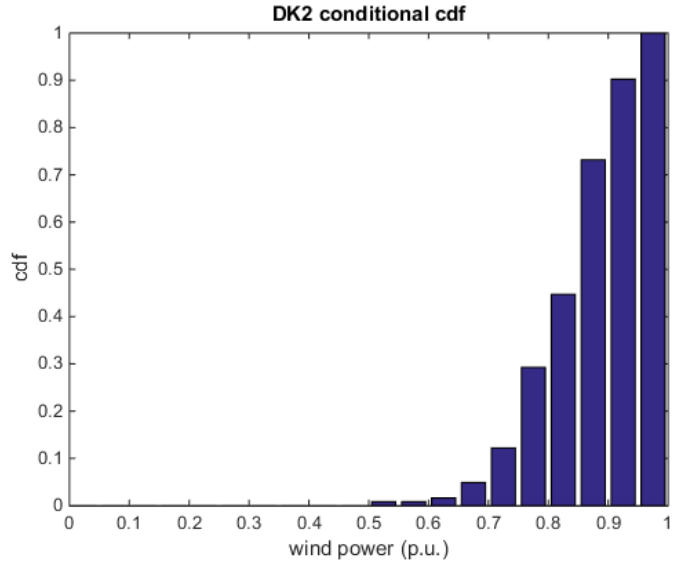


Figure 8: Conditional cdf for DK2 with DK1 wind power in the highest production bin

with an empirical estimate for the probability density function $f(x_i^t)$, Eq. (25) can be applied to produce a conditional density function.

Alternatively, if we treat DK1 as the first area and categorize all $t = 1 \ldots n$ historical wind power outputs x_1^t in a certain range to bin r_i, the conditional cumulative density function for the second area (e.g. DK2) can be estimated empirically using the following formula

$$F(x_2|x_1 \in r_i) = \frac{1}{m} \sum_{x_{2,j}} I(x_{2,j} \leq x_2), \tag{29}$$

where $x_{2,j} \in \{x_2^t | x_1^t \in r_i, t = 1 \ldots n\}$ and m is the number of historical outputs $x_{2,j}$ with x_1^t in the bin r_i. Then, Eq. (29) can be used together with the inverse transformation method to sample DK2 wind power production conditional to the sampled production in DK1. Figure 8 illustrates such a conditional cdf in the case with DK1 wind power in the highest production bin. The probability for low wind power production in DK2 is zero but high for almost full production, which indicates strong positive spatial correlation. The full procedure for obtaining probabilistic medium-term wind power scenarios is in Algorithm 2.

Two quantile forecasts with cumulative probabilities of 0.35 and 0.9 and the point forecast for DK1 and DK2 are presented in Figures 9(a) and 9(b). The point forecast is very high around

Algorithm 2: Procedure for generating medium-term wind power scenarios

1. Estimate the marginal probability distribution function of wind power production in each area from historical data. Model the spatio-temporal dependencies $c(\cdot)$ with the empirical copula density function in Eq. (28) or by fitting some parametric copula. Alternatively, use Eq. (29) to find the conditional cumulative density function for each area.

2. Sample the output of the first area in the beginning of the forecast horizon ($t = 1$) from its empirical cdf with the inverse transformation method. The random numbers used for drawing the sample should have a temporal correlation structure specified by the parameter ε.

3. Recursively sample the output of the ith area at $t = 1$ with the conditional cdf estimated in step 1, given $x_1^t, ...x_{i-1}^t$. The random numbers used for sampling may have different temporal correlation for each area.

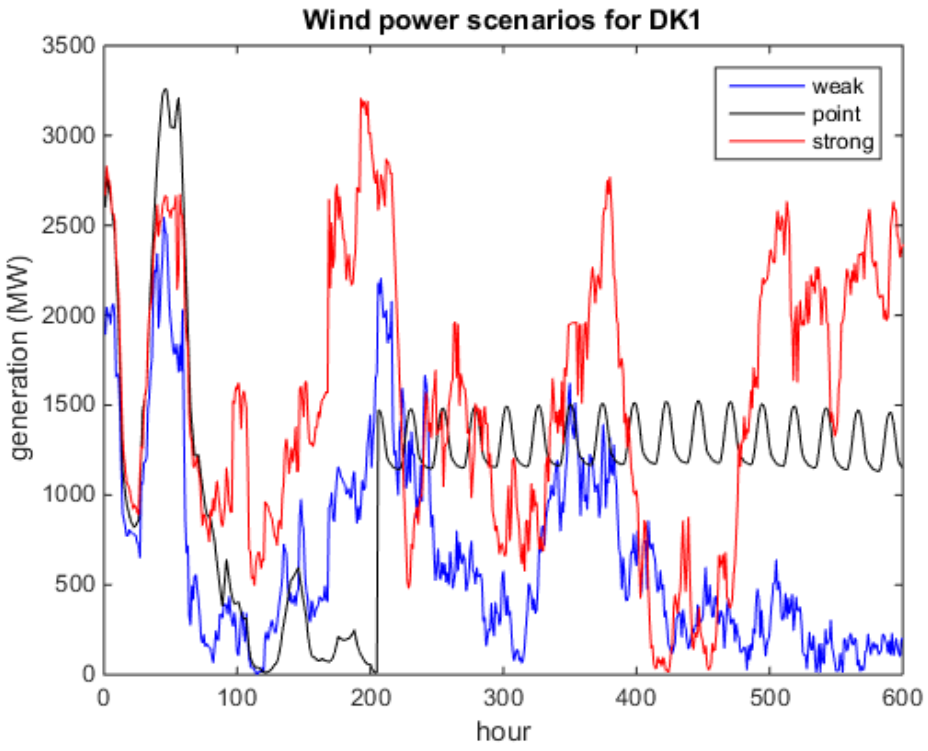
4. Repeat steps 2-3 until the final time period is reached.

hour 50 and then even the strong scenario is lower than the point forecast as the wind power forecast errors tend to be negative. In contrast, when the point forecast is very low, there is strong upside even in the weak scenario as the forecast errors are mainly positive. After the first 9-day period, both scenarios in both areas fluctuate considerably before a flatter period. Note that DK2 wind power is strongly positively correlated with DK1 as they are located close to each other.

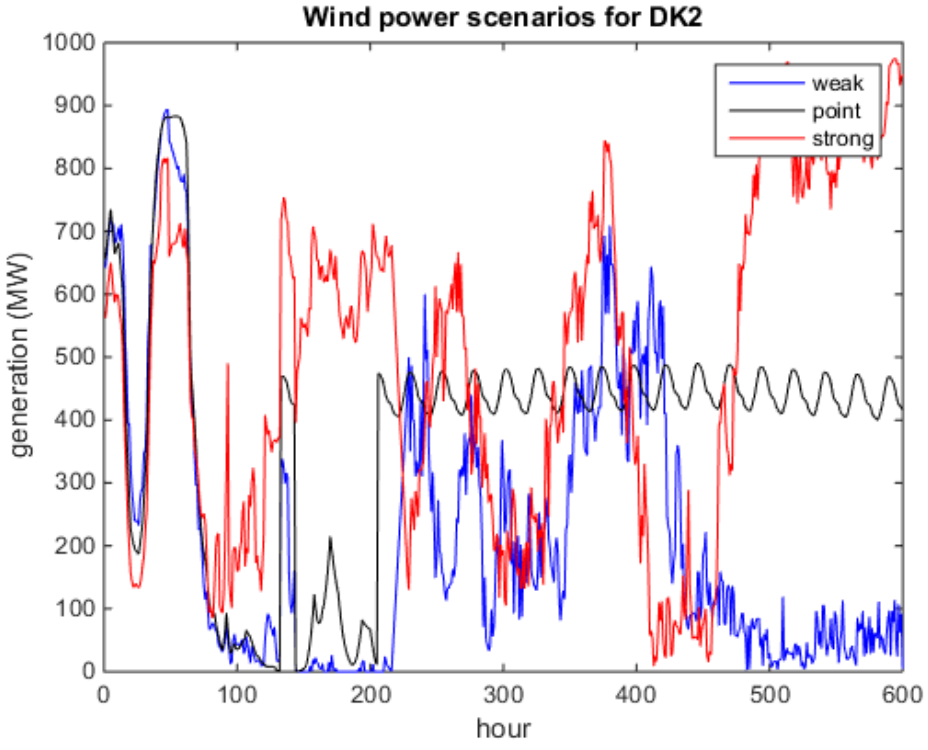
4.5 Demand scenarios

As for the case of wind power, the generation of demand scenarios is divided into two phases. For the first 9 days, a method equivalent to the one in Algorithm 1 is used to simulate demand forecast errors for different areas. A possible improvement of the method is to generate empirical distributions for working days and weekends separately but this was not implemented because the average forecast error was found to be merely 4% and 13% higher for working days than weekends in Finland and Sweden, respectively. Again, the temporal correlation of the forecast errors is adjusted with the ε parameter, and the calibration results are presented for Finland and Sweden in Figure 10. Similar to the calibration of the wind power model in Figure 7, the objective function value improves when ε starts to increase. The optimal value is now higher at approximately $\varepsilon = 170$ which indicates that there is more persistence in demand forecasts errors than those of wind power. This is explained by the high intermittency of wind power, which causes also the forecast errors to fluctuate more.

In the second phase, the point forecast f_{t+k} is based on an ensemble mean temperature forecast, historical average temperatures and demand. This point forecast is compared to the empirical probability distribution function $\hat{f}_h^d(x)$ of the realizations at the same time of the year during previous years (data set Y_h^d). The empirical distributions are generated separately for each hour h and working days and weekends d to model the differences between i) night time and peak hours ii) working days and weekend. However, the user should be careful in extending the data set Y_h^d far to the history, because long-term effects such as economic growth may make



(a) DK1



(b) DK2

Figure 9: Wind power scenarios for DK1 and DK2.

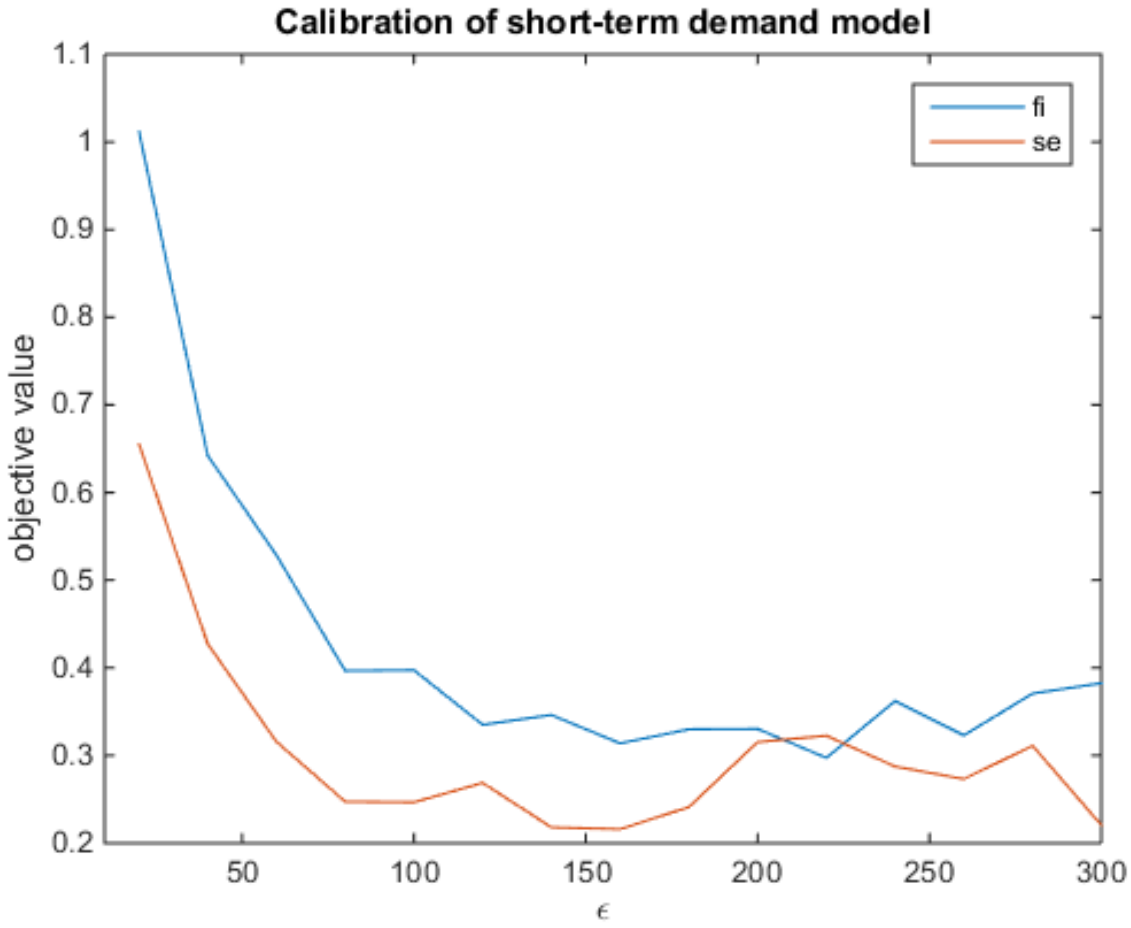


Figure 10: Calibration of the ε parameter for short-term demand forecast errors

the data inconsistent. Because of the stable demand development in recent years, our data Y_h^d starts from 2010. Then, separate cumulative distribution functions are computed for the probability mass on each side of the point forecast to determine the probability of higher and lower demand

$$\hat{F}^d_{h,\text{up}}(x) = \frac{1}{m_{\text{up}}} \sum_{y_i \in Y_h^d} I(y_i \leq x) \cdot I(y_i \geq f_{t+k}) \tag{30}$$

$$\hat{F}^d_{h,\text{down}}(x) = \frac{1}{m_{\text{down}}} \sum_{y_i \in Y_h^d} I(y_i \leq x) \cdot I(y_i \leq f_{t+k}), \tag{31}$$

where m_{up} and m_{down} are the number of historical observations above and below the point forecast, respectively. The method is visualised in Figure 11.

The scenarios are drawn from the upside and downside empirical cumulative distributions $\hat{F}^d_{h,\text{up}}$ and $\hat{F}^d_{h,\text{down}}$, respectively, by finding the inverse of the function at predetermined nominal proportions α_i. For example, demand on day d, hour h in upside scenario α_i is given by

$$x_h^{d,\alpha_i} = [\hat{F}^d_{h,\text{up}}]^{-1}(\alpha_i). \tag{32}$$

We note that the model could easily be modified by making nominal proportions α_i random variables. Successive α_i's should have strong temporal correlation because cold or warm spells tend to last several hours keeping the demand higher or lower for the same period, respectively. However, we keep α_i fixed to keep the strong demand scenarios continuously above weaker scenarios as we wish to explore the resulting interval of price scenarios.

In Figure 11, approximately 0.2 GW increase and 1 GW decrease have equal cumulative probability $\alpha = 0.5$. Therefore, there is more downside from the current forecast level than upside. This is likely to yield too high power price forecasts, which may result into too optimistic planning decisions. The methodology for generating medium-term wind power scenarios is summarised in Algorithm 3.

A major advantage of the proposed model is that the generated scenarios are based on the point forecast which has detailed seasonal information on demand. Also, holidays and bridging days can be taken into account by estimating separate distributions for them. On the other

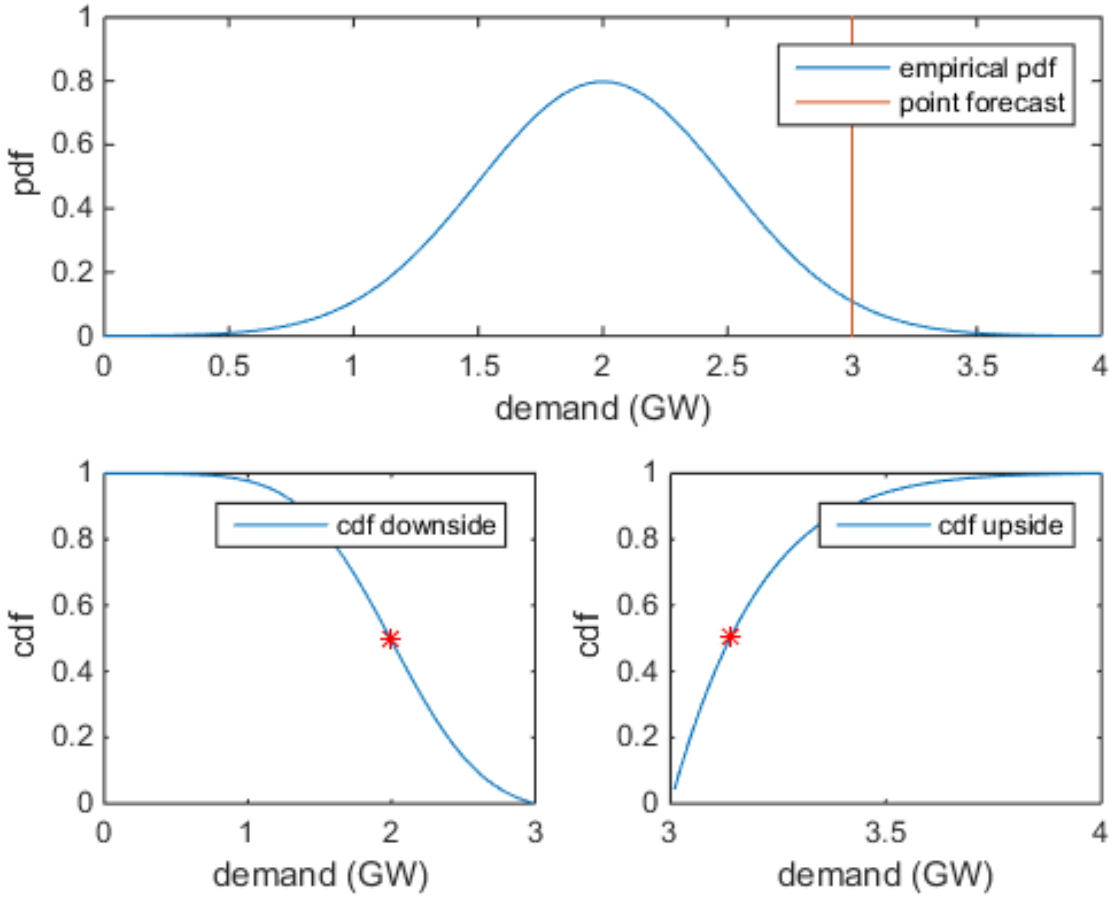


Figure 11: Separate empirical cumulative distributions for the probability mass below or above the current point forecast. The red dots indicate the sensitivity to up- and downside at the same cumulative probability 0.5

Algorithm 3: Procedure for generating medium-term demand scenarios

1. Estimate and store empirical probability density distributions $\hat{f}_h^d$ for each hour h for working days and weekend d separately using the historical data set Y_h^d.

2. Read the current point forecast f_{t+k} and determine which hour it is and whether it is a working day or weekend. Compare the point forecast to the corresponding empirical distribution $\hat{f}_h^d$. Compute the upside and downside cumulative distributions using Eqs. (30)-(31).

3. Given the upside and downside cumulative distributions $\hat{F}_{h,\text{up}}^d$ and $\hat{F}_{h,\text{down}}^d$, respectively, draw upside and and downside scenarios for hour h and day d using Eq. (32) with predetermined nominal proportions α_i.

4. Iterate steps 2-3 for all hours in the forecast horizon for each area.

hand, a drawback is that the model does not account for spatial correlation explicitly. Table 3 shows that the temporal correlation of two consecutive demand observations is approximately the same as in wind power, but the spatial correlation is higher. Therefore, when the medium-term model is used, one should use the same scenario for each area at the same time to include spatial correlation.

Two scenarios with cumulative probabilities of 0.2 and 0.8 in the short-term and 0.3 in the medium-term model for the whole Nord Pool are presented in Figure 12. The difference between the scenarios is relatively small in the beginning as the forecast errors are small. As the forecast errors are drawn using random numbers, it is temporarily possible that the strong scenario is lower than the base or even the weak scenario. However, as the forecast errors increase with time, the distance between the weak and strong scenario widens. Also, higher errors are caused by higher absolute level of the point forecast starting from hour 148. After 9 days, i.e., 216 hours, the upside is smaller than the downside as the point forecast is relatively high. Therefore, if the high demand does not realize, there is considerable downside risk.

	DK_t	DK_{t-1}	FI_t	FI_{t-1}	SE_t	SE_{t-1}
DK_t	1	0.96	0.76	0.76	0.83	0.81
DK_{t-1}		1	0.73	0.76	0.80	0.83
FI_t			1	0.98	0.95	0.92
FI_{t-1}				1	0.94	0.95
SE_t					1	0.98
SE_{t-1}						1

Table 3: Spatio-temporal correlations of Danish, Finnish and Swedish demand

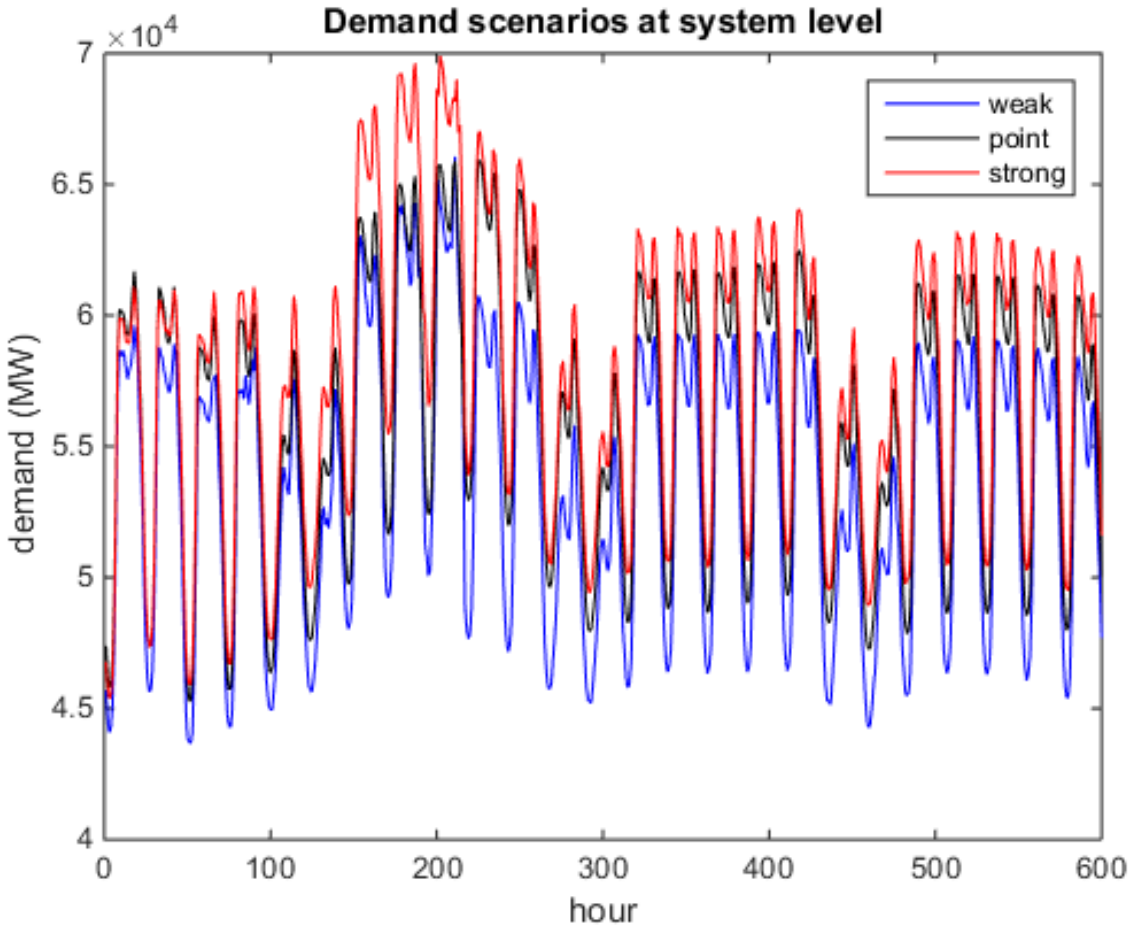


Figure 12: Demand scenarios for Nord Pool system level for selected quantiles.

4.6 Correlation between demand, wind power and cross-border capacities to Germany

Our models for wind power and demand scenario generation are not explicitly correlated. Figure 13 shows that the linear correlation coefficient of daily average wind power and demand can vary from negative to positive, depending on the month. It is widely known that low wind during winter is often associated with very cold temperatures and a high need for heating (negative correlation) while during summer, low wind may indicate warmer temperatures and less need for heating (positive correlation). The magnitude of the correlation is up to 0.6, which means that the impact of wind power and demand can be amplified considerably (negative correlation) or cancel themselves (positive correlation). However, linear correlation should be used with caution, because the relationship between wind power and demand is most likely nonlinear.

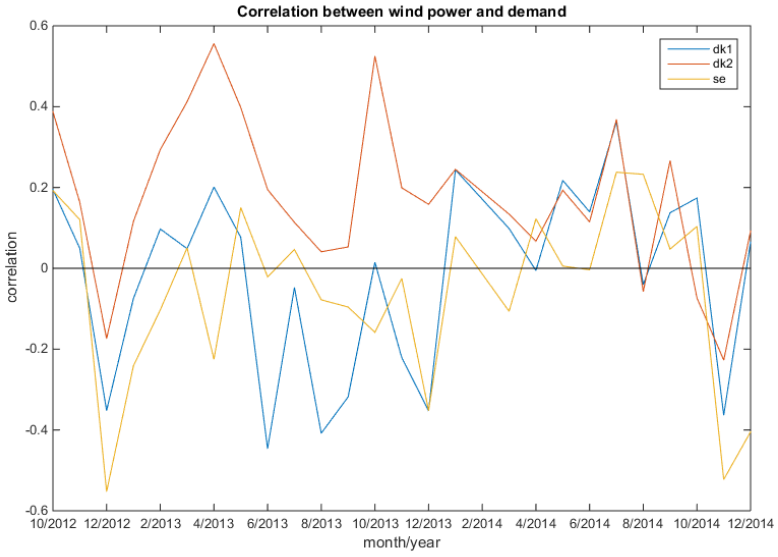


Figure 13: Correlation between daily average wind and demand

Similarly, wind power and demand forecast errors are found to exhibit both positive and negative correlation but the magnitude is weaker at a maximum of approximately 0.3. Weaker correlation is explained the stochastic nature of the forecast errors. However, as the forecast errors are situation dependent (see Figure 6) and wind power and demand forecast errors can

amplify each other, balancing market volumes could be modelled with our short-term wind power and demand scenarios.

To account for the complex correlation structure, the probabilities α_i in Eq. (32) could be random variables that are affected by the current wind power and vice versa. In practice, the exponential covariance function in Eq. (17) should be replaced with an empirical covariance estimate in Eq. (15), where the vector $\boldsymbol{Z}_t$ should contain both transformed wind power and demand errors. Alternatively, current demand could be modelled with Eq. (25) and conditioned with the recent wind power observations in adjacent areas. However, such models are left for future development, as they would require an integrated approach instead of creating separate models.

Moreover, in wind intensive Denmark and northern Germany, hourly wind power has an impact on hourly cross-border capacities, which, in turn, have an impact on power prices through limitations on export and import. The correlations of wind power and cross-border capacities are reported in Table 4. Wind power correlates negatively with the export capacity from the Nordics to Germany but positively with import through the line DE-DK1 and weakly negatively through the line DE-SE4, and, thus, export from Nordics is far more limited than import from Germany. However, appropriate modelling of these capacities would require a load flow model with security considerations; see Li et al. (2015) for an example on transmission line overloading in the case of correlated wind power and demand.

	DE-DK1$_h$	DK1-DE$_h$	DE-SE4$_h$	SE4-DE$_h$
DE wind$_h$	0.44	-0.48	-0.30	-0.31
DK1 wind$_h$	0.33	-0.43		
SE wind$_h$			-0.17	-0.14

Table 4: Correlations between wind power and cross-border capacities

5 Probabilistic accuracy of the demand and wind power scenarios

Statistical evaluation of weather dependent simulation data has been widely discussed in meteorological literature. Pinson et al. (2007b) present two criteria for assessing the probabilistic accuracy of the generated scenarios. First, the quantile forecasts should be reliable, i.e., the simulated quantiles should match the observed ones in Eq. (7). This requirement translates to the simulated forecasts being unbiased. The second requirement is the resolution of the probabilistic forecasts which stands for the ability to provide different distributions of the predictand in different situations (see Figure 6 for an example).

To verify the reliability of the proposed model for short-term wind power and demand scenario generation, we use a probability integral transform (PIT) histogram (Pinson et al., 2007b) which is built so that the proportion of the generated scenarios between two consecutive quantiles should equal to the difference between the nominal probabilities of these two quantiles

$$\frac{1}{N}\sum_{j=1}^{N} I(q^{\alpha_i} \leq \hat{x}_j \leq q^{\alpha_{i+1}}) = \alpha_{i+1} - \alpha_i \qquad i = 1 \ldots n, \tag{33}$$

where $\hat{x}_i$ are the simulated realizations, q^{α_i} the observed quantiles and α_i the associated nominal proportions. Figures 14 and 15 show the PIT histogram with 20 quantiles for DK1 wind power and FI demand forecast errors with 9-day look-ahead time and 4-week evaluation data set from January 2015. Both scenarios are slightly biased downwards as the nominal probabilities for small quantiles are higher than 0.05, and, thus, cases where the forecast is too high are overrepresented.

The reliability of the medium-term models can be verified with the PIT histogram as well. Now, we generate 30000 instances of 21-day wind power generation and categorize each sim-

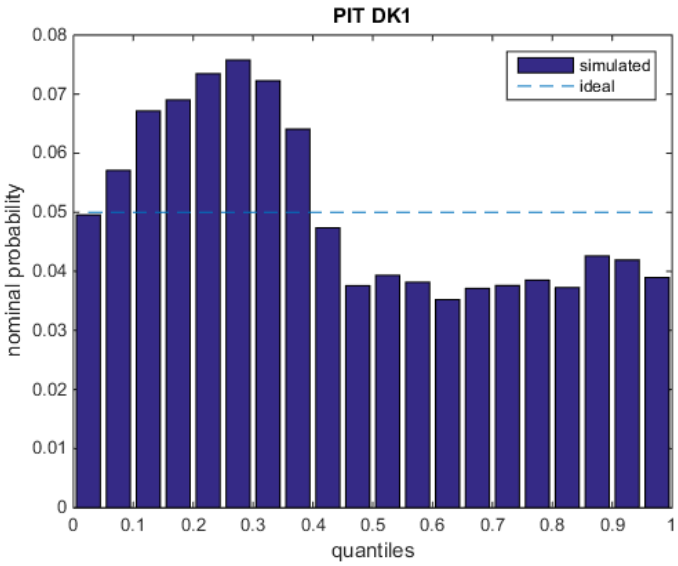


Figure 14: PIT histogram of DK1 short-term wind power scenarios

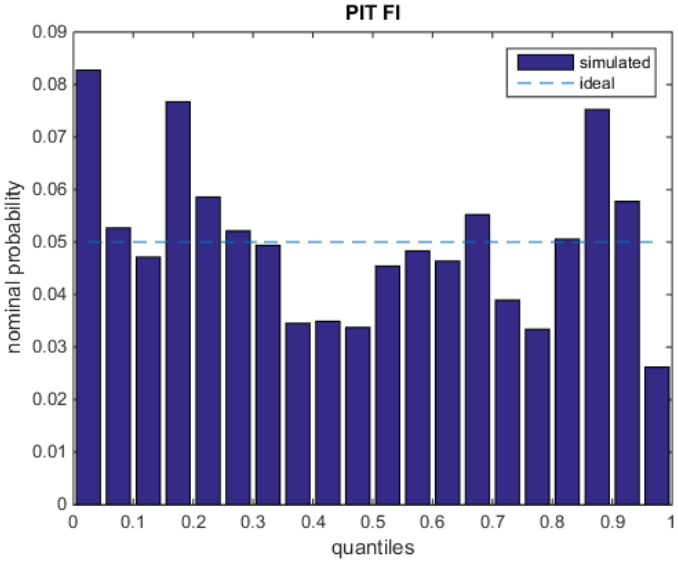


Figure 15: PIT histogram of FI short-term demand scenarios

ulated value according to Eq. (33). PIT histograms for DK1 and DK2 wind power scenarios in Figure 16 show some upward bias for the highest quantiles. A possible explanation is that the increase in wind power capacity over time causes large outputs to be sampled although such values are not present in the historical data. In our data set starting from the end of 2012, the increase in DK1 wind power capacity was approximately 35% and in DK2 capacity 10%. A smaller data set could resolve the issue but then the estimation of the medium-term model would become less accurate. Already now, the method may get stuck into extreme values because, as Figure 8 shows, there is very low probability for low wind power. A similar bias is present both in DK1 and DK2 because they are spatially correlated.

However, the PIT histogram for Finnish medium-term demand scenarios in Figure 17 shows very good overall reliability. Only some of highest quantiles are slightly overrepresented meaning that the scenarios are slightly too high. The reason for better performance is that the demand model uses data starting from 2010 and the change in demand has been lower than that for wind power. Moreover, demand is easier to model statistically.

Table 5 shows that the instantaneous spatial correlations of the evaluation data set for medium-term wind power are close to the realised ones in Table 2. Correct spatial correlations of wind power are important for modelling area prices as they have not only an impact on the areawise supply but also on the transmission flows between the Nordic countries. Moreover, the measured slight upward bias in the scenarios is not amplified by too high spatial correlation which would cause too high wind simultaneous wind power in all areas.

	$DK1_t$	$DK2_t$	SE_t
$DK1_t$	1	0.81	0.61
$DK2_t$		1	0.59
SE_t			1

Table 5: Simulated instantaneous spatial correlations

Moreover, all methods studied in this thesis should fulfill the requirement for resolution by definition. Short-term models for wind power and demand estimate the empirical cumu-

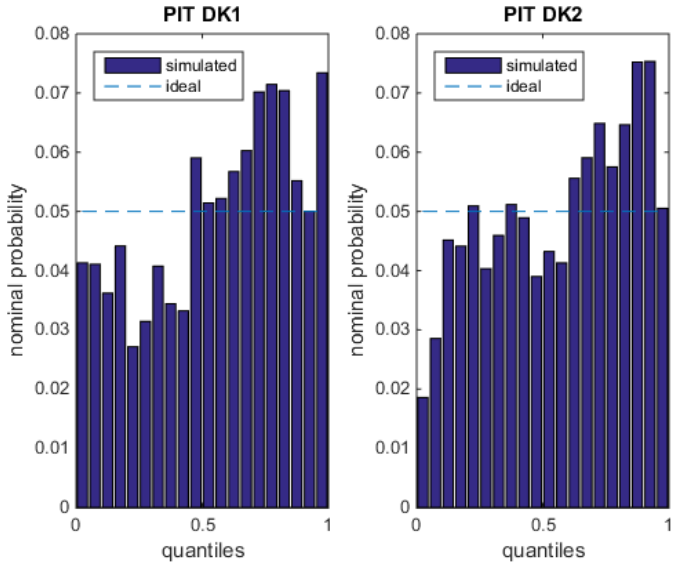


Figure 16: PIT histogram of DK1 and DK2 medium-term wind power scenarios

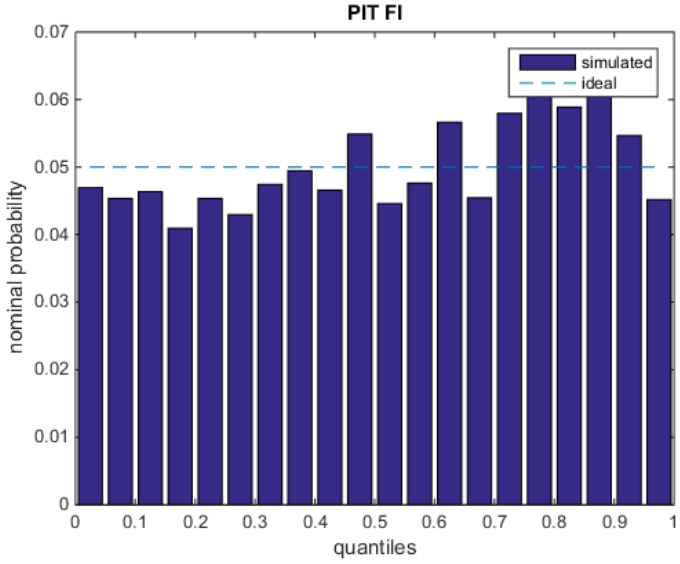


Figure 17: PIT histogram of FI medium-term demand scenarios

lative distribution function for each forecast condition and look-ahead time separately, while the medium-term model for demand compares the current point forecast to recent historical realizations and the one for wind power is calibrated to the first differences of realised wind power to capture temporal correlation correctly, i.e., extended periods of very low or high wind are statistically unlikely.

Following Pinson et al. (2007b), we measure the resolution of the quantile forecasts by computing the average interval between two quantiles $\hat{q}_{t+k|t}^{(1-\beta/2)}$ and $\hat{q}_{t+k|t}^{\beta/2}$

$$\delta_k^\beta = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{q}_{t+k|t}^{(1-\beta/2)} - \hat{q}_{t+k|t}^{\beta/2}\right), \tag{34}$$

where β determines the quantile, k is the look-ahead time, t the forecast time and N the number of forecasts. Figure 18 shows the δ-plots of the short-term wind power and demand models for several look-ahead times and quantiles. In both models, the resolution weakens with look-ahead time, but the model for demand has tighter intervals. Consequently, demand forecasts involve less uncertainty than wind power forecasts although demand is significantly higher than wind power in absolute terms. Due to the look-ahead resolution of one day, the intervals tend to widen every 24 hour. This could be mitigated by increasing the look-ahead resolution, but, on the other hand, it would decrease the accuracy of the empirical cumulative distribution functions by reducing the available data for each distribution. In the demand scenarios, night time have higher resolution than peak hours, whereas wind power scenarios show no clear seasonality.

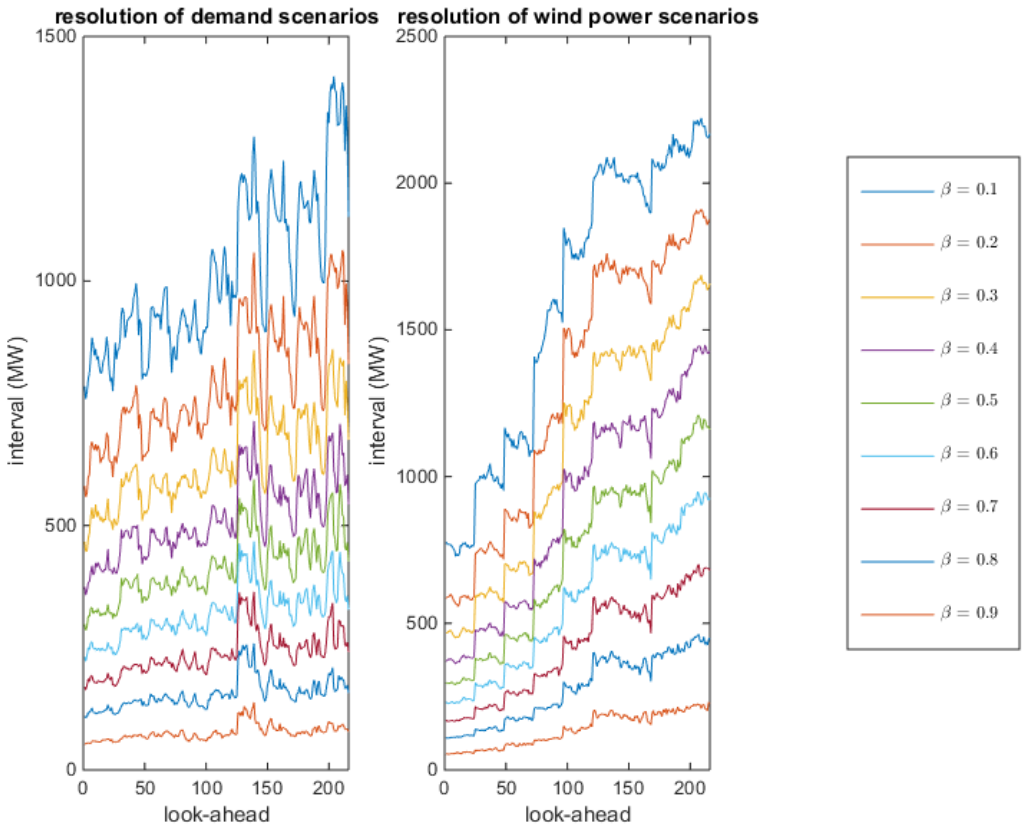


Figure 18: Resolution of short-term wind power and demand models

6 Power price forecasts based on demand and wind power scenarios

We use the wind power and demand scenarios from Sections 4.4 and 4.5, respectively, as an input to the day-ahead market model in Eqs. (1)-(4) to study their marginal impacts on system and area prices as well as volatility. Demand scenarios have a direct impact on the variable $d_{n,t}$ in Eq. (2) but wind power scenarios are reflected in the parallel shifts of the cost functions $s_{n,t}(\cdot)$. Because the size of the German power system is large compared to the capacity of the transmission lines between Germany and the Nordic countries, the cross-border exchange was optimized against a fixed German price forecast, i.e., German prices were assumed to be insensitive to the exchange. Often, German day-ahead prices differ from Nordic prices by several euros so even if the sensitivity was modelled, the changes in exports should be very small. To account for the correlations between the demand and wind power scenarios, their compound impact is analyzed. Finally, we simulate with different levels of cross-border capacity between Germany and the Nordic countries to find insight in price converge and the propagation of high peak and low night time prices.

First, Figure 19 shows how the wind scenarios in Figures 9(a) and 9(b) affect Nord Pool system price forecast in a weak demand scenario. The forecast with point forecasts as input is referred to as base forecast. When the wind power point forecast is very high near hour 50 and the both wind power scenarios are below it, the weak and strong price forecast scenarios have higher night time prices. Around hour 200, the strong price forecast scenario has considerably lower peak prices as the wind power point forecast is very low and there is considerable upside to it. At the end of the forecast horizon, there is a possibility for low night time prices in the strong wind power scenario because wind power is very high both in DK1 and DK2 in

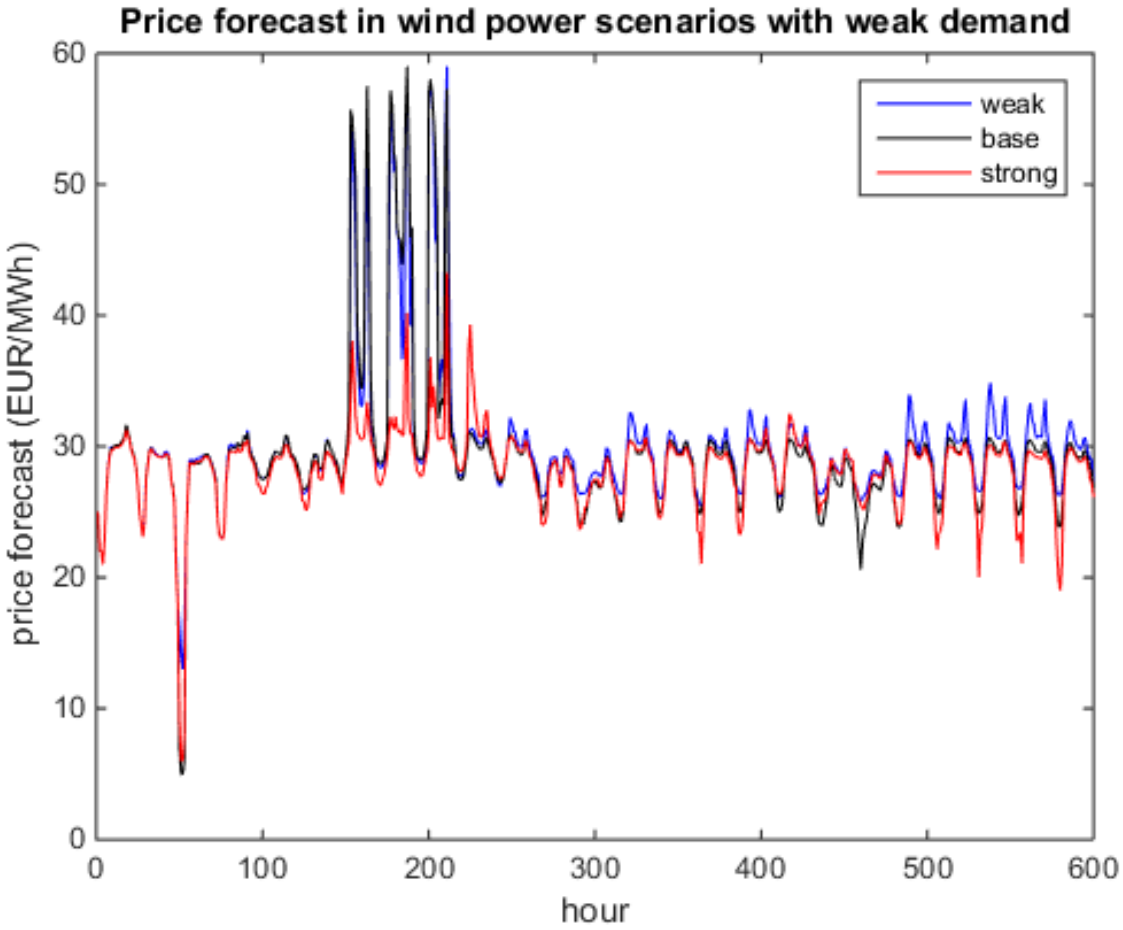


Figure 19: Price forecasts in different wind scenarios with weak demand

Figures 9(a) and 9(b), respectively. On the other hand, the weak wind power scenario shows a possibility for higher peak prices because DK1 and DK2 wind power are simultaneously low.

These examples show that the situation dependent information of wind power forecast errors is transferred to the power price forecast. The differences between power price scenarios are considerable and as such they would result into different planning decisions. Moreover, the possibility for low night time prices would be underestimated if the strong spatio-temporal correlation of wind power was not modelled.

Next, we plot the different wind power scenarios with the point and strong demand levels in Figures 20 and 21, respectively. In the former case, peak prices rise to 60 €/MWh level and in the latter as high as 200 €/MWh, which is an unlikely peak price level but which occurs every

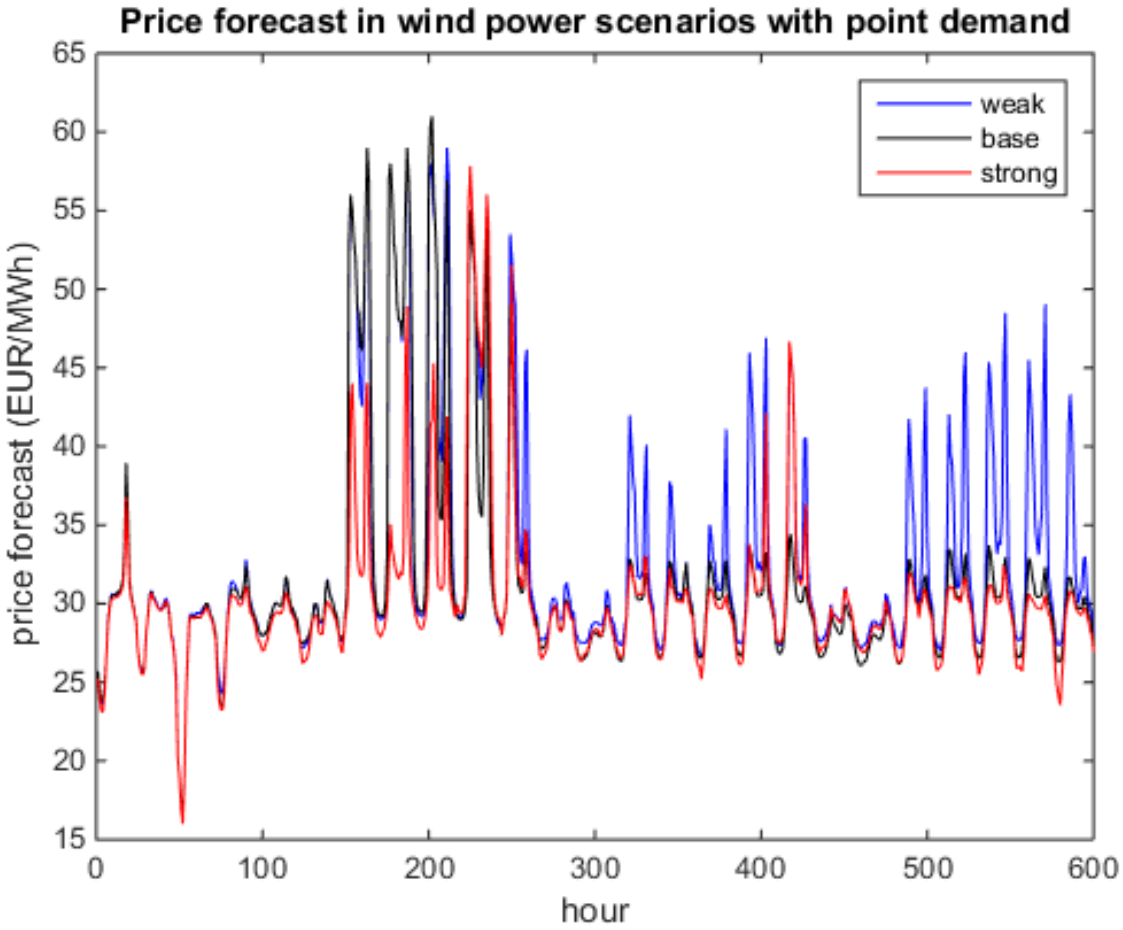


Figure 20: Price forecasts in different wind scenarios with point demand

now and then. In both demand scenarios, night time prices do not drop below 15 €/MWh level in any wind scenario, contrary to scenario with weak demand, indicating that such a low night time prices are unlikely. As Figure 4 shows, system price sensitivity is higher in the upper steep part of the supply curve than in the lower steep part, and, thus, peak prices shoot up fast whereas night time prices do not crash as much. Also, increasing demand has more impact on the average price than increasing wind power because the differences between the scenarios are larger for demand than for wind power.

Figure 22 shows a wind power-demand scenario matrix in terms of volatility in different areas. Volatility increases significantly at the system level and in all areas when demand increases. However, the impact of wind power is inconclusive in this forecast horizon because

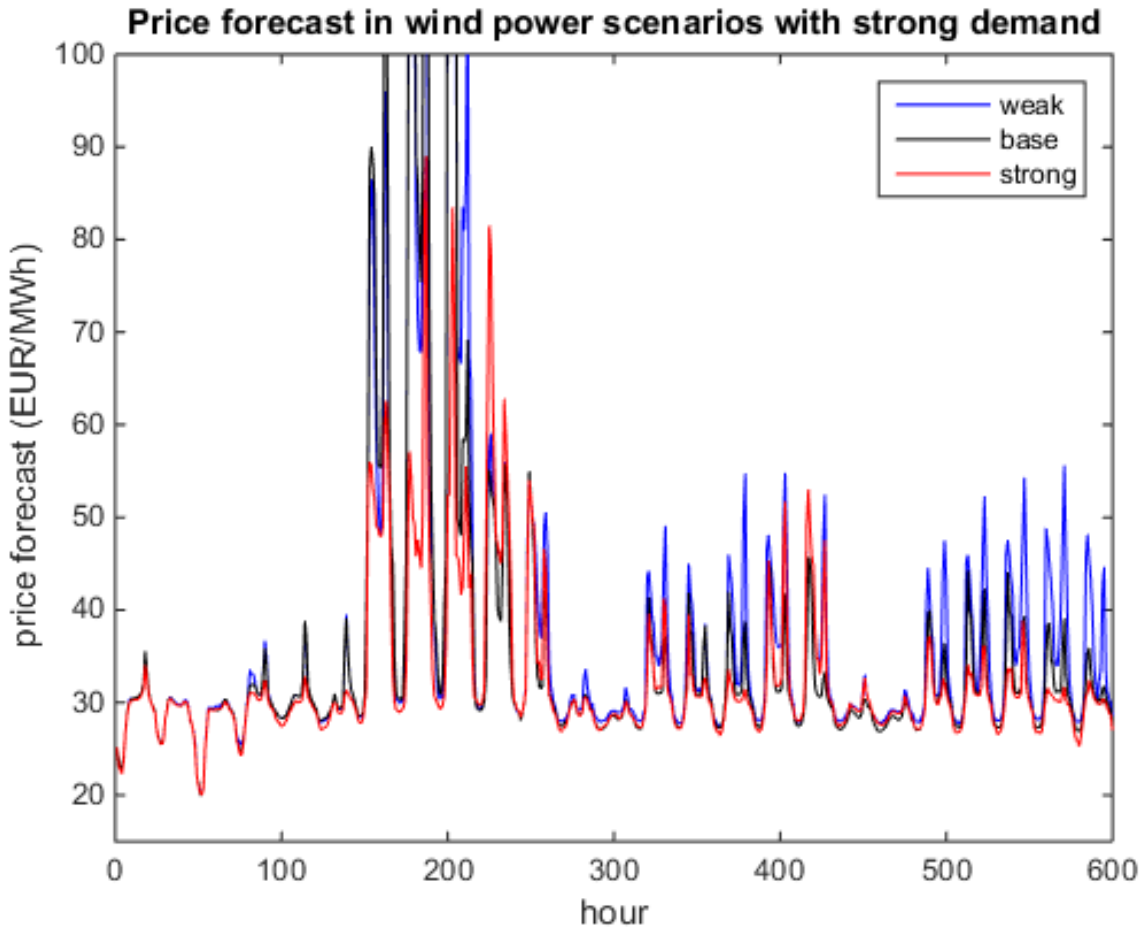


Figure 21: Price forecasts in different wind scenarios with strong demand

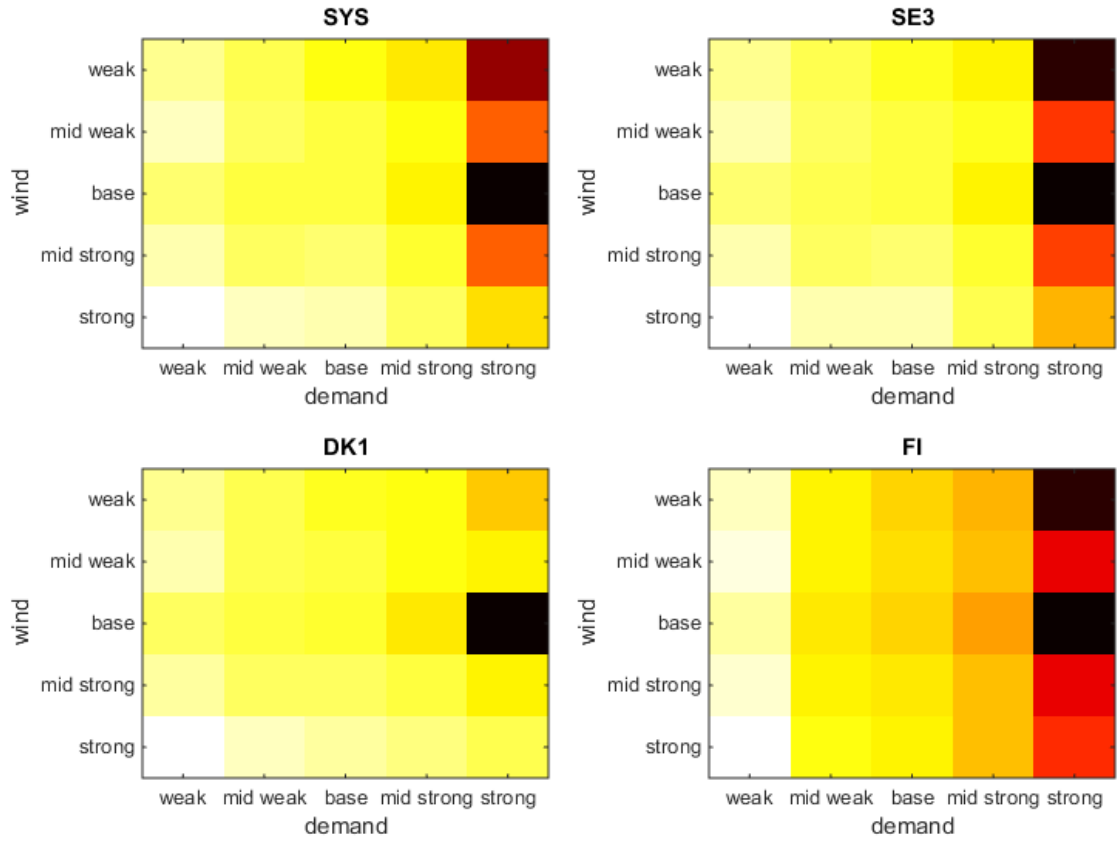


Figure 22: Wind power-demand scenario matrix of weekly volatility. Darker colours indicate higher volatility and lighter lower.

high wind power causes variation in consecutive hourly prices but low wind power increases the volatility by increasing peak prices considerably. DK1, SE3 and SYS are most sensitive to wind power but FI not due to the lack of wind power and import capacity. FI, on the other hand, is the most sensitive to changes to demand as supply is scarce.

In Table 6, demand and wind power point forecasts are used but the cross-border capacity between the Nordic countries and Germany changes. Here, the weak capacity scenario refers to a case in which 10% of installed capacity is available and 90% in the strong scenario. In this example, Germany has higher weekly prices than the adjacent Nordic areas. When the cross-border capacity increases, area prices converge. Regardless of the area price convergence, Table 7 shows that volatility increases in all areas.

The price convergence and higher volatility are caused by the changes in net exchange, which is shown in Figure 23. Nordic countries import from Germany during the morning off-peak hours but export otherwise, and, thus, the Nordic area prices decrease during off-peak but increase in the peak. However, Finland is mainly affected by the lower off-peak prices because Finnish peak prices often decouple from Sweden due to the dependency on imports. Consequently, the change in Finnish average price and price volatility is lower than those for DK1, SE3 and SYS.

	Capacity weak					Capacity strong				
Area / Week	DE	DK1	FI	SE3	SYS	DE	DK1	FI	SE3	SYS
1	39.5	35.8	36.7	35.8	35.8	39.5	37.7	38.2	37.7	37.5
2	31.4	29.0	33.0	29.1	29.1	31.4	29.5	32.9	29.5	29.8
3	29.5	28.9	34.0	29.0	29.0	29.5	29.2	33.8	29.2	29.5

Table 6: System and area prices in two different cross-border capacity scenarios with point demand and wind power forecast. Note that German price is fixed. All figures are in €/MWh.

	Capacity weak					Capacity strong				
Area / Week	DE	DK1	FI	SE3	SYS	DE	DK1	FI	SE3	SYS
1	12.2	10.9	11.1	10.9	10.9	12.2	11.1	11.1	11.1	10.9
2	9.8	1.4	7.4	1.4	1.5	9.8	2.6	7.5	2.6	3.0
3	9.4	1.4	8.3	1.4	1.5	9.4	2.4	8.5	2.4	2.9

Table 7: System and area price volatility in two different cross-border capacity scenarios with point demand and wind power forecast. Note that German price is fixed. All figures are in €/MWh.

In Section 4.6 the correlation between wind power and demand scenarios was found to be negative or positive depending on the month. In the case with negative correlation, the

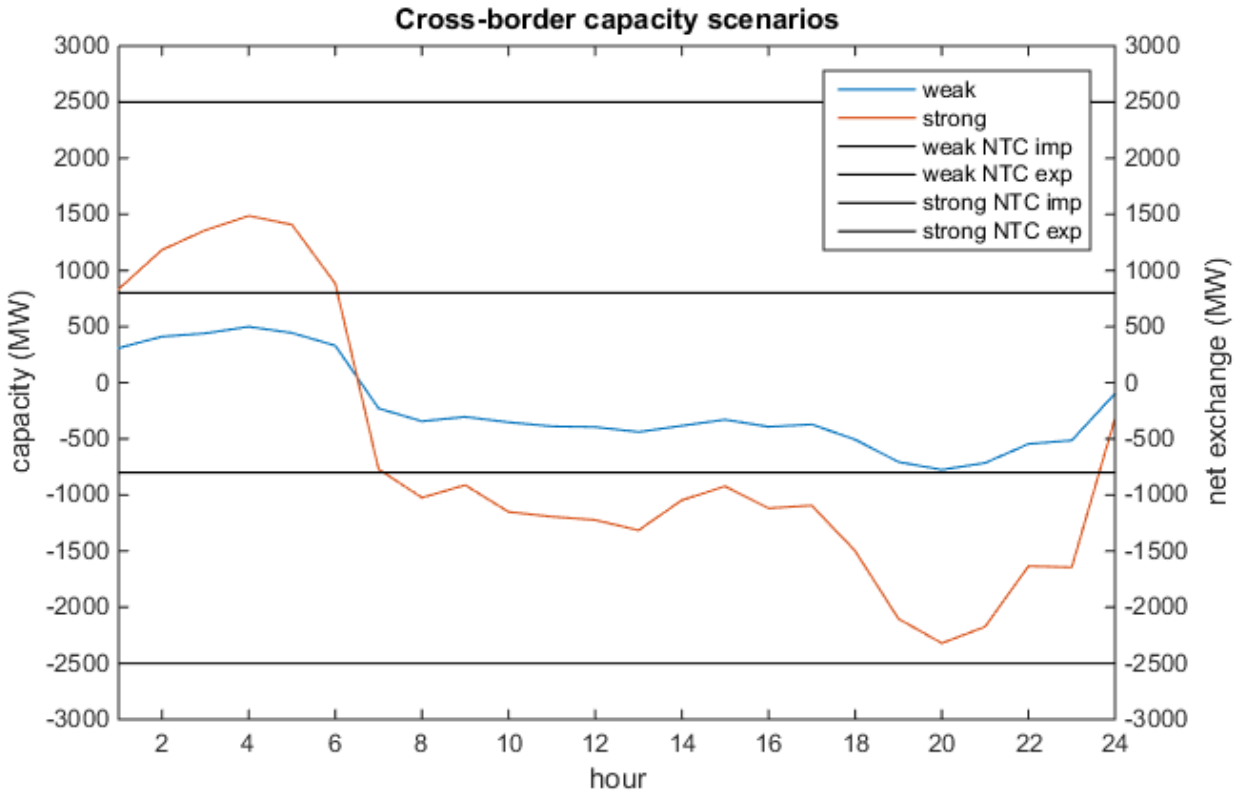


Figure 23: Maximum export (exp) and import (imp) capacities and average hourly net exchange in different capacity scenarios

wind power and demand scenarios amplify each other. Figures 19-21 show that opposite wind power and demand scenarios (strong wind, weak demand, for example) lead to either very low night time or high peaks prices with high overall price volatility. When the correlation is positive, same wind power and demand scenarios cancel out each other to some extent. Unfortunately, our methods do not give the joint probability of the scenarios so the likelihood of the aforementioned scenarios cannot be evaluated.

Finally, we explore the difference in hourly system prices in a scenario i) with strong demand and weak wind and ii) with weak demand and strong wind, when cross-border capacity drops from 90% availability to 10% and increases from 10% to 90%, respectively. The former scenario corresponds to a situation, in which the cross-border capacities would be limited from the current levels and the latter scenario corresponds to improved availability especially in the case of high power generation. When cross-border capacity becomes limited in the former scenario, very high peak prices between hours 100-200 in Figure 21 increase, because imports from Germany drop. On average, peak price level increases by 0.3 €/MWh and night-time hours by 0.25 €/MWh. When the cross-border capacity increases and the Nordic daytime prices are below German prices, as is the case in the latter scenario, then higher export to Germany results into 0.7 €/MWh higher peak prices. However, night time prices drop by 0.4 €/MWh when low prices from Germany spread out to the Nordic countries. Consequently, Nordic and German prices are found to converge if the availability of transmission capacity is improved.

7 Conclusion

Variable wind power and demand cause high uncertainty to power prices and pose challenges to the operation of the system. Consequently, the value of a single point forecast diminishes and probabilistic forecasts need to be employed. Such forecasts provide the decision maker information about short-term price peaks and drops, which represent positive and negative risks for the producers in the day-ahead market, respectively. Furthermore, the scenarios present situation dependent and quantitative information about the possible upside and downside in medium-term prices, which make them valuable for production allocation purposes and financial trading.

Therefore, we have developed models for short- and medium-term wind power and demand scenario generation, which consider spatio-temporal correlations and the sensitivity of the point forecast to up- and downside in each situation. Our statistical evaluation of the reliability of the models shows that the short-term models and the medium-term model for demand perform relatively well. However, in the medium-term model for wind power there is some upward bias, which is most likely caused by the increase in wind power capacity and the inaccuracies in the empirical distributions. The observed spatial correlations are shown to be modelled precisely.

The models use relevant historical data to predict future developments in each situation to fulfill the requirement for resolution. The adequate resolution of the short-term models for wind power and demand was tested and it was found that it deteriorates over time as expected. The resolution of wind power forecasts was shown to be lower than that of demand because wind power fluctuates more. The resolution of both models could be improved by increasing the look-ahead resolution.

We feed the generated scenarios to a price calculation model which imitates the social wel-

fare maximization conducted by the electricity exchanges. In periods with negative correlation between wind power and demand, very high or, on the contrary, very low prices can be achieved. On the other hand, when the correlation is positive, changes in wind power and demand offset each other to some extent. We have also observed price convergence between the Nordic countries and Germany if the installed transmission capacity is fully available. Currently, wind power correlates negatively with the export capacity to Germany but positively with import capacity to Nordics, and, thus, low night-time and high peak hour prices may propagate from Germany to the Nordic countries and increase volatility. Also, our model gives insight into area price differences such as the one between Finland and Sweden and shows how the low night time prices caused by high wind power generation in Denmark, Sweden and Germany may propagate to Finland, too.

The models are separated for wind power and demand. This allows us to study their statistical properties and impact on power prices individually. However, modelling their mutual relationship becomes difficult although there exists a correlation between the wind power and demand scenarios. Consequently, the next step could be an integrated model combining wind power and demand using copulas, for example, to establish the relationship explicitly and yield more realistic results. Alternatively, the covariance matrix Σ used in sampling could be represented as a block matrix with a wind and demand submatrices. Such an integrated model would also provide correct probabilities for each wind power-demand scenario, which could be used to form a probability weighted base forecast. The integrated model could also be extended to cover transmission capacities and, thus, provide an even stronger coupling between different inputs. Moreover, phenomena such was wind power ramps and generation cut-offs could also be modelled.

However, implementing the integrated model would require a large historical dataset of historical demand and wind speed forecasts and their realizations so that the high-dimensional copula functions could be calibrated accurately enough. Instead of wind power generation, wind speed data should be used to remove the impact of the increase in wind power capacity. However, as the generated scenario tree grows exponentially, one should find a balance between

the coverage of the model and the computational time required for simulating power prices without drawing the model unsuitable for operational use.

An automatized version of the process introduced in this thesis was tested in daily four-week price forecasting and it was found to be light enough for operative scenario generation. Outside the day-ahead market, a natural application for our short-term models is probabilistic modelling of intraday and balancing market volumes conditioned on the current forecast level and market situation. This is supported by the correlation between wind power and demand forecast errors. Such a model could help reduce the risk of not being dispatched in the intraday markets and, at the same time, maximize dispatchable volumes in scenarios with high intraday prices.

From the power system perspective, we see that the extreme events caused high intermittency and forecast errors of renewable generation could be mitigated by i) incentivising flexible generation and demand response, ii) expanding the transmission system, and iii) integrating European balancing markets to enable imbalance netting between the different areas. Also, extending the area of the integrated electricity market could mitigate the risks posed by the spatial correlation of weather dependent factors. Ultimately, however, increased distributed generation may motivate consumers to leave the transmission grid. Hence, developing storage technologies and regional optimisation of distributed assets (see e.g. Kraning et al., 2013) should be promoted to mitigate the intermittency of renewable generation.

References

Baringo, L. and Conejo, A. (2011). Wind power investment within a market environment. *Applied Energy*, 88:3239–3247.

Baringo, L. and Conejo, A. (2013). Correlated wind-power production and electric load scenarios for investment decisions. *Applied Energy*, 101:475–482.

Bludszuweit, H., Domínguez-Navarro, J. A., and Llombart, A. (2008). Statistical analysis of wind power forecast error. *IEEE Transactions on Power Systems*, 23(3):983–991.

Boomsma, T. K., Juul, N., and Fleten, S.-E. (2014). Bidding in sequential electricity markets: The Nordic case. *European Journal of Operational Research*, 238:797–809.

Bundesnetzagentur (2014). Monitoringbericht 2014. `http://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Bundesnetzagentur/Publikationen/Berichte/2014/Monitoringbericht_2014_BF.pdf?__blob=publicationFile&v=4`. [Online; accessed 9 January 2015].

Charpentier, A., Fermanian, J.-D., and Scaillet, O. (2006). The estimation of copulas: Theory and practice. In Rank, J., editor, *Copulas: From Theory to Application in Finance*. Risk books.

EEX (2014). Power trading results in September 2014 - intraday volume grows 46% year-on-year. `http://www.epexspot.com/en/press-media/press/details/press/Power_Trading_Results_in_September_2014_-_Intraday_volume_grows_46_year-on-year`. [Online; accessed 13 October 2014].

Egerer, J., Kunz, F., and von Hirschhausen, C. (2013). Development scenarios for the North and Baltic seas grid - a welfare economic analysis. *Utilities Policy*, 27:123–134.

ENTSO-E (2014). Ten-year network development plan 2014. Map of projects long-term. https://www.entsoe.eu/Documents/TYNDP%20documents/TYNDP%202014/140710_TYNDP_2014_Projects_of_European_relevance_LT_2019-2030.pdf. [Online; accessed 16 September 2014].

Eurelectric (2013). Power statistics trends 2013 edition. http://www.eurelectric.org/media/113667/power_statistics_and__trends_2013-2013-2710-0001-01-e.pdf. [Online; accessed 10 April 2015].

Farahmand, H. and Doorman, G. (2012). Balancing market integration in the Northern European continent. *Applied Energy*, 96:316–326.

Feng, Y., Gade, D., Ryan, S., Watson, J.-P., Wets, R.-B., and Woodruff, D. (2013). A new approximation method for generating day-ahead load scenarios. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5.

Foley, A., Gallachóir, B. , Hur, J., Baldick, R., and McKeogh, E. (2010). A strategic review of electricity systems models. *Energy*, 35:4522–4530.

Gabriel, S. A., Conejo, A. J., Hobbs, B. F., Fuller, J. D., and Siddiqui, S. S. (2013). *Complementarity Modeling in Energy Markets*. Springer, New York.

Gabriel, S. A. and Leuthold, F. U. (2010). Solving discretely-constrained MPEC problems with applications in electric power markets. *Energy Economics*, 32(1):3–14.

GAMS (2015). General algebraic modeling system. http://www.gams.com/. [Online; accessed 13 October 2014].

Geman, H. and Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *Journal of Business*, 79(3):1225–1261.

Grothe, O. and Schnieders, J. (2011). Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, 39(9):4742–4754.

Growe-Kuska, N., Heitsch, H., and Romisch, W. (2003). Scenario reduction and scenario tree construction for power management problems. In *Power Tech Conference Proceedings, 2003 IEEE Bologna*, volume 3, pages 1–7.

Hobbs, B. F. (2001). Linear complementarity models of Nash–Cournot competition in bilateral and POOLCO power markets. *IEEE Transactions on Power Systems*, 16(2):194–202.

Hodge, B.-M. and Milligan, M. (2011). Wind power forecasting error distributions over multiple timescales. In *Power and Energy Society General Meeting, 2011 IEEE*, pages 1–8.

Huovila, S. (2003). Short-term forecasting of power demand in the Nord Pool market. Master's thesis, Lappeenranta University of Technology.

Jaehnert, S. and Doorman, G. L. (2012). Assessing the benefits of regulating power market integration in Northern Europe. *Electrical Power and Energy Systems*, 43:70–79.

Jaehnert, S., Wolfgang, O., Farahmand, H., Völler, S., and Huertas-Hernando, D. (2013). Transmission expansion planning in Northern Europe in 2030 — methodology and analyses. *Energy Policy*, 61:125–139.

Jordan, M. I. (2010). Monte Carlo sampling. `http://www.cs.berkeley.edu/~jordan/courses/260-spring10/lectures/lecture17.pdf`. [Online; accessed 3 February 2015].

Ketterer, J. C. (2014). The impact of wind power generation on the electricity price in Germany. *Energy Economics*, 44:270–280.

Klæboe, G., Eriksrud, A., and Fleten, S.-E. (2015). Benchmarking time series based forecasting models for electricity balancing market prices. *Energy Systems*, 6(1):43–61.

Kraning, M., Chu, E., Lavaei, J., and Boyd, S. (2013). Dynamic network energy management via proximal message passing. *Foundations and Trends in Optimization*, 1(2):70–122.

Kristiansen, T. (2014). A time series spot price forecast model for the Nord Pool market. *Electrical Power and Energy Systems*, 61:20–26.

Kunz, F. (2013). Improving congestion management: How to facilitate the integration of renewable generation in Germany. *Energy Journal*, 34(4):55–78.

Kännö, J. (2013). A short-term price forecast model for the Nordic electricity markets. Master's thesis, Aalto University School of Science.

Leuthold, F. U., Weigt, H., and von Hirschhausen, C. (2012). A large-scale spatial optimization model of the European electricity market. *Networks Spatial Economics*, 12(1):75–107.

Li, X., Zhang, X., Wu, L., Lu, P., and Zhang, S. (2015). Transmission line overload risk assessment for power systems with wind and load-power generation correlation. *IEEE Transactions on Smart Grid*, 6(3):1233–1242.

Ma, X.-Y., Sun, Y.-Z., and Fang, H.-L. (2013). Scenario generation of wind power based on statistical uncertainty and variability. *IEEE Transactions on Sustainable Energy*, 4(4):894–904.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., and Kjeldsen, T. (2014). Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology*, 519 D:3634–3650.

Mauritzen, J. (2013). Dead battery? Wind power, the spot market, and hydropower interaction in the Nordic electricity market. *The Energy Journal*, 34(1):103–123.

McSharry, P. E., Bouwman, S., and Bloemhof, G. (2005). Probabilistic forecasts of the magnitude and timing of peak electricity demand. *IEEE Transactions on Power Systems*, 20(2):1166–1172.

Neuhoff, K., Barquin, J., Bialek, J. W., Boyd, R., Dent, C. J., Echavarren, F., Grau, T., von Hirschhausen, C., Hobbs, B. F., Kunz, F., Nabe, C., Papaefthymiou, G., Weber, C., and Weigt, H. (2013). Renewable electric energy integration: Quantifying the value of design of markets for international transmission capacity. *Energy Economics*, 40:760–772.

Nielsen, H., Nielsen, T., Madsen, H., Giebel, G., Badger, J., Landbergt, L., Sattler, K., Voulund, L., and Tofting, J. (2006). From wind ensembles to probabilistic information about future wind power production – results from an actual application. In *International Conference on Probabilistic Methods Applied to Power Systems*, pages 1–8.

Nord Pool Spot (2014). An effective European power market through price coupling of regions. `http://www.nordpoolspot.com/How-does-it-work/European-Integration/Price-coupling-of-regions/`. [Online; accessed 15 September 2014].

Nord Pool Spot (2015). Price calculation principles. `http://nordpoolspot.com/TAS/Day-ahead-market-Elspot/Price-calculation/Price-calculation-principles/`. [Online; accessed 26 January 2015].

Nowotarski, J. and Weron, R. (2014). Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Computational Statistics*, 30(1):1–13.

OffshoreGrid (2014). OffshoreGrid project. `http://www.offshoregrid.eu/`. [Online; accessed 16 September 2014].

Pinson, P., Chevallier, C., and Kariniotakis, G. (2007a). Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, 22(3):1148–1156.

Pinson, P. and Kariniotakis, G. (2010). Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems*, 25(4):1845–1856.

Pinson, P., Nielsen, H. A., Møller, J. K., Madsen, H., and Kariniotakis, G. N. (2007b). Non-parametric probabilistic forecasts of wind power: Required properties and evaluation. *Wind Energy*, 10:497–516.

Pinson, P., Papaefthymiou, G., Klöckl, B., and Nielsen, H. A. (2011). Generation of statistical scenarios of short-term wind power production. *Applied Energy*, 88(2):432–441.

Qin, Z., Li, W., and Xiong, X. (2013). Incorporating multiple correlations among wind speeds, photovoltaic powers and bus loads in composite system reliability evaluation. *Applied Energy*, 110:285–294.

Ruiz, C., Conejo, A. J., and Bertsimas, D. (2013a). Revealing rival marginal offer prices via inverse optimization. *IEEE Transactions on Power Systems*, 28(3):3056–3064.

Ruiz, C., Conejo, A. J., Fuller, J. D., Gabriel, S. A., and Hobbs, B. F. (2013b). A tutorial review of complementarity models for decision-making in energy markets. *EURO Journal on Decision Processes*, 2(1-2):91–120.

Spiecker, S., Vogel, P., and Weber, C. (2013). Evaluating interconnector investments in the North European electricity system considering fluctuating wind power penetration. *Energy Economics*, 37:114–127.

Taylor, J. W. and Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, 19:57–70.

The European Wind Energy Association (2014). Wind in power. 2013 European statistics. `http://www.ewea.org/fileadmin/files/library/publications/statistics/EWEA_Annual_Statistics_2013.pdf`. [Online; accessed 09 January 2015].

Valizadeh Zaghi, H. and Lotfifard, S. (2015). Spatiotemporal modeling of wind generation for optimal energy storage sizing. *IEEE Transactions on Sustainable Energy*, 6(1):113–121.

Vehviläinen, I. and Pyykkönen, T. (2005). Stochastic factor model for electricity spot price—the case of the Nordic market. *Energy Economics*, 27(2):351–367.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081.

Würzburg, K., Labandeira, X., and Linares, P. (2013). Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria. *Energy Economics*, 40(S1):S159–S171.

Zhang, N., Kang, C., Xu, Q., Jiang, C., Chen, Z., and Liu, J. (2013). Modelling and simulating the spatio-temporal correlations of clustered wind power using copula. *Journal of Electrical Engineering Technology*, 8(6):1615–1625.

Zhou, M., Yan, Z., Ni, Y., Li, G., and Nie, Y. (2006). Electricity price forecasting with confidence-interval estimation through an extended ARIMA approach. *IEEE Proceedings on Generation, Transmission and Distribution*, 153(2):187–195.

Zugno, M., Pinson, P., and Madsen, H. (2013). Impact of wind power generation on European cross-border power flows. *IEEE Transactions on Power Systems*, 28:3566–3575.